Methods of Computational Neuroscience

Part 4 of 6

Marine Biological Laboratory - Woods Hole, MA
August 2009

MCN 2009 Lecture Notes & Papers
Reader – Part 4 of 6

Week 2 (Aug. 10-14)

Day 1 (Aug 10)

Bard Ermentrout

- chapt1.pdf
- A very comprehensive tutorial on bifurcations for neuro
- A dated tutrial for XPP
- A very quick look at using xpp to solve the HH eqns
- A link to the xpp homepage

ml.ode Morris-Lecar ODE
hhh.ode Hodgkin-Huxley

Installation of XPP on your computer:

- Using XPP on windows

Adrienne Fairhall

- mcn2009Fairhall_2.pdf
- hongAgueraFairhallNC07.pdf
- lundstromFamulareFairhall.pdf
- agueraFairhallBialek.pdf
- hongLundstromFairhallPloSCB08.pdf
- lundstromNC08.pdf

Day 2 (Aug 11)

Dan Johnston

Day 3

Michael Hausser

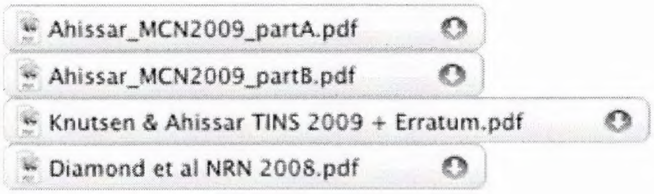
- Hausser_slides.pdf
- Sjostrom_Hausser_Neuron06.pdf
- Hausser&London2005review.pdf

MCN 2009 Lecture Notes & Papers

Reader – Part 3 of 6

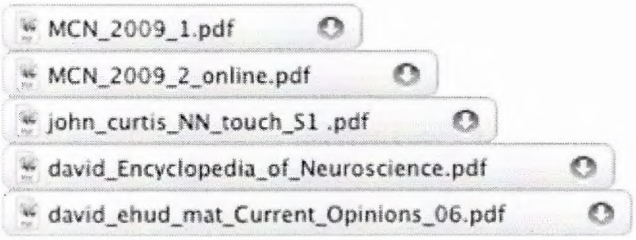
Day 4 (Aug. 6)

Ehud Ahissar



- ○ Ahissar_MCN2009_partA.pdf
- ○ Ahissar_MCN2009_partB.pdf
- ○ Knutsen & Ahissar TINS 2009 + Erratum.pdf
- ○ Diamond et al NRN 2008.pdf

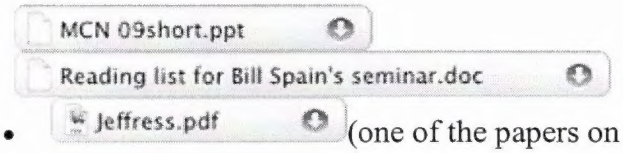
David Kleinfeld (website)



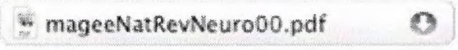
- MCN_2009_1.pdf
- MCN_2009_2_online.pdf
- john_curtis_NN_touch_S1 .pdf
- david_Encyclopedia_of_Neuroscience.pdf
- david_ehud_mat_Current_Opinions_06.pdf

Day 5 (Aug 7)

Bill Spain



- MCN 09short.ppt
- Reading list for Bill Spain's seminar.doc
    - Jeffress.pdf (one of the papers on the reading list that is not available online)

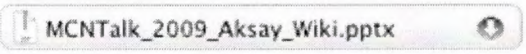
- 
mageeNatRevNeuro00.pdf

Bartlett Mel

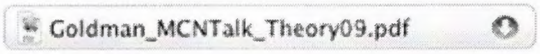
-
Mel_MCN_lecture.pdf

Day 4 (August 13)

Emre Aksay

- 
MCNTalk_2009_Aksay_Wiki.pptx

Mark Goldman

- 
Goldman_MCNTalk_Theory09.pdf

MCN 2009 Lecture Notes & Papers
Reader – Part 5 of 6

Day 5 (August 14)

Nancy Kopell

Sara A. Solla

- MBL09_SOLLA.ppt
- PRL04.pdf
- Chaos07.pdf

Week 3 (Aug. 17-21)

Day 1 (Aug. 17)

Karel Svoboda

- Svoboda_MCN.pdf

Larry Abbott

- AbbottLect1.pdf
- force5.pdf
- Bi, G.-q. & Poo, M.M. Synaptic modification by correlated activity: Hebb's postulate revisited. Annu. Rev. Neurosci. 24,139-166.

- Caporale N. & Dan, Y. Spike Timing-Dependent Plasticity: A Hebbian LearningRule. Annu. Rev. Neurosci. 31, 25-46 (2008).
- Dan Y. & Poo M.M. Spike timing-dependent plasticity: from synapse to perception. Physiol Rev. 86, 1033-1048 (2006).

Day 2 (August 18)

Sara Solla

▶ 120:34

Haim Sompolinsky:

- Learning spike based rules-THE TEMPOTRON
- guetig_sompolinsky06_full.pdf
- gutig_sompolinsky_plos09.pdf

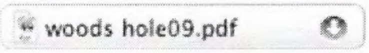
Larry Abbott's second talk: Chaotic networks

- AbbottTalk2.pdf
    - Jaeger H., and Haas H. (2004). Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. Science 304, 78-80.

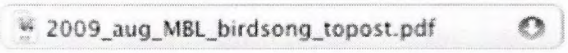
MCN 2009 Lecture Notes & Papers Reader – Part 6 of 6

Day 3 (August 19)

Bence Olveczky

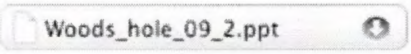
- 
  woods hole09.pdf

Ila Fiete

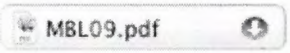
- 
  2009_aug_MBL_birdsong_topost.pdf

Day 4 (August 20)

Reza Shadmehr

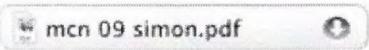
- 
  Woods_hole_09_2.ppt

Emo Todorov

- 
  MBL09.pdf

Day 5 (August 21)

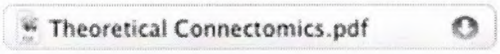
Simon Laughlin

- 
  mcn 09 simon.pdf

Week 4 (Aug. 24-28)

Day 1 (Aug. 24)

Mitya Chklovskii

- 
  Theoretical Connectomics.pdf

Problem Sets

Week 1

Simulated Neurons

Problem Set 1.zip

See the papers from day 1 (Fairhall et al., 2006 and Pillow et al., 2008)

Information theory

ProblemSet-InformationTheory.pdf

ProblemSet-FisherInfoCorrelatedNeurons-MCN2009.pdf (updated 2009.08.11)

Solutions

ProblemSetSolution-InformationTheory-MCN2009.pdf (updated 2009.08.11)

ProblemSetSolution-FisherInfoCorrelatedNeurons-MCN2009.pdf (updated 2009.08.11)

Spectral analysis

Link to Chronux download website

Kleinfeld_Mitra_CSHL_2009_revised..pdf

matlab_for mcn_2990.m

example_data.mat

whisking.mat

090606_PAS_noStim.mat

Week 2

Working with XPP

NOTE: These problems are to be done at your discretion, and will not be presented.

- tcurrent.ode Exploring T-current and bifurcations
- ping.ode A simple model for PING

These are text files. The instructions are at each individual file after the 'Done' statement.

NEURON tutorial and exercises

- NEURON Tutorial ppt.pdf
- NEURON short exercises.pdf
    - Complete these exercises to become familiar with NEURON, as well as gaining a better understanding of dendritic dynamics.
- simple_layerVpyramidal.zip
- Webpage for downloading NEURON to your computer
    - Follow the instructions for your OS of choice. At the computer stations in the lab we use the MacOS-X 10.4 version
    - For Mac laptops, you HAVE to install the XCode package (size is about 1GB). Please ask the TA's for it, before the tutorial on Wednesday.

NEURON mini projects

- NEURON mini projects.pdf
- Rall_MCN_Assignment.pdf
- Katz et al 2009 - Neuron.pdf
- Kampa & Stuart 2006 - JNeurosci.pdf

This version of the code contains an additional .hoc file for use with the second mini project. Open "exerciseS2.hoc" to access it.
mbl0813.zip

for problem 3:
Calcium spikes in basal dendrites (Kampa and Stuart 2006) [modelDB]

Attractor network exercises

Mark Goldman will discuss your solutions with you on Monday afternoon.
MCN_RecurrentNetsExercise.pdf

MCN 2009 Lecture Notes & Papers
Reader – Part 1 of 6

Week 1 (Aug. 3-7)

Day 1 (Aug. 3) - Reader – Part 1 of 6

Adrienne Fairhall

- mcn09FairhallSlides.pdf
- Fairhall2006.pdf (Covariance analysis in retina)
- Pillow2008.pdf (GLM analysis in retina)
- Fairhall-etal-01.pdf (Multiple timescale adaptation in the fly)
- lundstromNN08.pdf (Fractional differentiation in cortical neurons)
- MaravallPLOS07.pdf (Gain control in barrel cortex)
- ruderman1997.pdf (Review of natural scene statistics)

Bill Bialek (website)

- bialek_02.pdf Les Houches: *Thinking about the brain*
- The ideas about optimal motion estimation go back to a paper with Marc Potters, which can be found here: http://www.princeton.edu/~wbialek/our_papers/potters+bialek_94.pdf
- The material on photon counting in vision is covered in a set of lecture notes which can be found at

  http://www.princeton.edu/~wbialek/PHY562.html

MCN 2009 Lecture Notes & Papers
Reader – Part 2 of 6

Day 2 (Aug. 4) - Reader – Part 2 of 6

Jonathan Pillow

- pillow_slides09_MaxLi_GLM.pdf (lecture slides)

Michael Berry (website)

- MCN 2009 Populations.ppt
- Recording All Cells.pdf
- Redundancy in Retinal Code.pdf

Rava da Silveira

- TheOrchestralBrain-WithAnnotations.pdf

Day 3 (Aug. 5)

Roozbeh Kiani

- Kiani_CategorizationInIT.pdf
- kianiJNP07.pdf
- kianiNeuron08.pdf

Jim DiCarlo

- DiCarloWHslides.pdf

Elad Schneidman

- Schneidman-MBL-MCN-2009.pdf
- Schneidman_2006_Nature_Weak pairwise correlations imply strongly.pdf
- Schneidman_PRL_2003_Network information and connected correlations.pdf
- Tkacik+al_06.pdf

1 Differential Equations

Bard Ermentrout and John Rinzel

1.1 What Is a Differential Equation?

All processes in neuroscience change over time and it is thus necessary to model these processes as dynamic systems, that is, systems whose current and future states can be determined from their past. The most common and simplest types of dynamic models come in the form of differential equations in which the rate of change of some variable, say $u(t)$, is given as a function of u and other quantities. For example,

$$\frac{du}{dt} = -u/\tau + I(t). \tag{1.1}$$

This simple differential equation compactly embodies the rule for prediction of the system's entire future. It states that the instantaneous rate of change, increase or decrease, of $u(t)$ is the sum of a memory loss term with rate $-u(t)/\tau$ and an input rate $I(t)$. Many dynamic processes can be described by differential equations as instantaneous rules in this way:

$$\frac{du}{dt} = \text{change per unit time} = \text{sum of production rates} - \text{sum of removal rates}. \tag{1.2}$$

As a familiar example, suppose we are modeling the time course of a neuron's membrane potential $V(t)$. If c is the capacitance of the neuron's membrane, then the balance law for growth or decay of a charge on the membrane is

$$c\frac{dV}{dt} = I_{\text{ext}} - \sum I_{\text{inward}}(V, t) - \sum I_{\text{outward}}(V, t), \tag{1.3}$$

where we have grouped the different inward and outward currents that flow through the membrane's ionic channels. This is a differential equation for $V(t)$. It is called an *ordinary* differential equation (ODE) because there is only one independent variable: time, t.

More generally, we may combine the production and removal terms in a net input rate function $G(u,t)$. Here, by including u as as an argument in G, we allow that G would generally have contributions that depend on $u(t)$ itself, and perhaps nonlinearly:

$$\frac{du}{dt} = G[u(t),t]. \tag{1.4}$$

A complete formulation for predicting the future, say, from some particular starting time t_0 requires that we also specify the state at $t = t_0$, $u(t_0)$. Prediction corresponds to solving the differential equation given the initial condition.

How do we solve the differential equation (1.4) for $u(t)$? How do we obtain the prediction for the future of u? One can imagine doing this computationally by marching forward in time, taking small steps in time of size Δt. That is, we make the approximation and replace the differentials dt and du by differences Δt and $\Delta u = u(t + \Delta t) - u(t)$. Then equation (1.4) becomes

$$u(t + \Delta t) = u(t) + \Delta t G[u(t),t]. \tag{1.5}$$

This straightforward recipe is easy to implement and is guaranteed to work if Δt is small enough and if the differential equation is reasonable (see Mascagni and Sherman, 1998). That is, speaking more rigorously, under reasonable assumptions about the function $G(u,t)$—namely, the derivative of G with respect to u is continuous and G is a continuous function of time—this equation guarantees that there exists a unique solution to the differential equation (1.4) for all times t close enough to t_0.

Why do we care about this mathematical arcana? Since most differential equations that are written for neurons do not have a nice closed-form solution [that is, we cannot write an expression for the solution in terms of known functions like $\sin(t)$], we must generally use a computer to find solutions. The point of existence and uniqueness is that they tell you that there is a solution which, with care, can be well approximated by the computer. Later on we will see that there are some types of differential equations that do not have any solutions unless the parameters are carefully chosen. (Some ODE problems specify more conditions than there are ODEs and so will have solutions only for special values of the parameters. We illustrate this in the section on partial differential equations later in this chapter with the example of finding the speed and profile of a propagating impulse for a Hodgkin-Huxley-like cable equation.)

The formulation for a single unknown quantity, u, can be generalized to systems of differential equations, which we can write as

$$\begin{aligned}\frac{dx_1}{dt} &= G_1(x_1, \dots, x_n, t) \\ \frac{dx_2}{dt} &= G_2(x_1, \dots, x_n, t) \\ \cdots &= \cdots \\ \frac{dx_n}{dt} &= G_n(x_1, \dots, x_n, t).\end{aligned} \tag{1.6}$$

These equations are all evaluated in parallel. That is, given the values of all the n variables at the current time t, their values at $t + \Delta t$ are approximately

$$x_j(t + \Delta t) \approx x_j(t) + \Delta t G_j. \tag{1.7}$$

That is, the functions G_j specify the rates of change of each of the variables x_j and how they depend on the present values of those variables. We call an equation like (1.6) a *system of first-order differential equations*. For example, if the $x_j(t)$ terms represent, say, concentrations of chemicals involved in a set of chemical reactions, then the functions G_j are the rates of change for each chemical species and depend on the specific reactions and their reaction rates. In neuroscience applications, the variables x_j usually are the potentials of neurons, the fractions of open channels, and concentrations of ionic species. We remark that the time scales of all these interacting species can vary over many orders of magnitude. Yet, once their values are specified at a single point in time, we can know them for any other points in time for which they are defined.

Many physical phenomena involve more than just the rates of change. For example, a mass m attached to a spring with a spring constant k (ignoring friction) obeys Newton's equation:

$$\text{mass} \times \text{acceleration} = \text{force}, \tag{1.8}$$

which in terms of differential equations is

$$m \frac{d^2x}{dt^2} = -kx. \tag{1.9}$$

Here the displacement of the mass (relative to rest), $x(t)$, is determined by its second derivative rather than just its rate of change. This is a second-order differential equation, and in order to specify $x(t)$ we need to know both the position, $x(t_0)$, and the velocity, $dx/dt(t_0)$, at the same time t_0. If we introduce a new variable, the velocity, defined as $q(t) = dx/dt$, then we can write the equations for the mass-spring system as

$$\frac{dx}{dt} = q$$
$$\frac{dq}{dt} = -(k/m)x. \tag{1.10}$$

Here we use the fact that $dq/dt = d^2x/dt^2$. We have reduced the second-order equation to a pair of first-order equations. In many cases of interest, systems of nth-order differential equations can be transformed into systems of first-order equations by introducing intermediate variables. The converse is not true: we cannot write most systems of first-order equations as a single nth-order system. For this reason, we will stick with systems of first-order differential equations in this chapter. (Most simulators of differential equations require that you write them as systems of first-order equations, so it is good to practice this.)

If the functions $G_1, \ldots, G_n$ are independent of the variable t, we call the system of equations *autonomous*. Otherwise this is a nonautonomous differential equation. In general, autonomous equations are easier to analyze. The behavior of solutions to ODEs can be as simple as convergence to a constant in time, or they can vary in an oscillatory fashion or even become chaotic.

Another class of models consists of *delay-differential* equations, in which the evolution rule depends not only on the current state but also on the system's state some finite time in the past. This might be useful, say, as a simplified way of accounting for propagation or synaptic delays between neurons in a network model. A simple example is the delayed negative feedback model:

$$\frac{du}{dt} = -u(t) + \frac{a}{1 + bu(t - t_d)^p}. \tag{1.11}$$

While the formulation is simple-looking, such models can have quite complex behaviors. The solution at time t depends, not on just the present time, but rather on a whole interval of the past. In order to compute behavior at a time t_0, the behavior of $u(t)$ must be known for all $t_0 - t_d \le t \le t_0$. The analysis and simulation of delay equations require special methods and numerical techniques beyond the scope of this chapter. For descriptions of some dynamic properties and analysis of delay equations in biology, see Macdonald (1989).

In this chapter we will highlight a few properties of differential equations and concepts for understanding them. If you wish more elaboration, but without overburdening mathematical detail, we recommend the books by Edelstein-Keshet (2005), Izhikevich (2006), and Strogatz (2001), and the appendices in Fall et al. (2005).

1.2 Linear Systems

If the functions G_j are linear with respect to the variables, x_j, then we call the system *linear*. For example, equation (1.1) is a one-variable (or first-order) linear differential equation. If $I(t)$ is a constant, $\bar{I}$, the solution is a simple exponential:

$$u(t) = \bar{I}\tau + (u_0 - \bar{I}\tau)\exp[-(t - t_0)/\tau], \tag{1.12}$$

where $u_0 = u(t_0)$. Linear systems arise naturally as descriptions when nonlinearities are weak and negligible and, more generally, for considerations of stability and sensitivity of nonlinear systems in the neighborhood of an equilibrium operating point.

In the general case of n-variables it is convenient to write the system (1.6) in vector-matrix form:

$$\frac{d\boldsymbol{X}}{dt} = \boldsymbol{A}(t)\boldsymbol{X}(t) + \boldsymbol{f}(t). \tag{1.13}$$

Here $\boldsymbol{A}(t)$ is an $n \times n$ matrix, $\boldsymbol{X}(t)$ is the vector of variables $(x_1, \dots, x_n)$, and $f(t)$ is the n-component vector of t-dependent input functions. For example, for the spring-mass system (1.10) in vector-matrix form, we would have the vector $\boldsymbol{X} = (x, q)$ and the matrix:

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ -k/m & 0 \end{pmatrix}. \tag{1.14}$$

If the matrix entries and the function, $f(t)$, are continuous with respect to time, then there exists a unique solution to the linear equation with $\boldsymbol{X}(t_0) = \boldsymbol{X}_0$ for all time. This is much better than we can hope for in the nonlinear case. [The reader should verify that the solution to $dx/dt = x^2$, with $x(0) = 1$ is $1/(1 - t)$, which exists only for $t < 1$.] We say that a linear system is *homogeneous* if $f(t) = 0$. Linear homogeneous autonomous equations are the only general class of equations for which an explicit solution can always be written. If the entries of A are constant in time, the solution to

$$\frac{d\boldsymbol{X}}{dt} = \boldsymbol{A}\boldsymbol{X} \tag{1.15}$$

is a sum of n exponentials:

$$\boldsymbol{X}(t) = c_1\boldsymbol{V}_1 e^{\lambda_1 t} + \cdots + c_n \boldsymbol{V}_n e^{\lambda_n t}. \tag{1.16}$$

The c_j are arbitrary constants and the pairs $\boldsymbol{V}_j$, λ_j comprise the n “modes” of the system, determined solely by the matrix $\boldsymbol{A}$. Since each exponential term of this sum must satisfy (1.15), we have (by substituting $V_j e^{\lambda_j t}$ into equation 1.15) that

$$\lambda_j \boldsymbol{V}_j = \boldsymbol{A}\boldsymbol{V}_j. \tag{1.17}$$

These scalar-vector pairs are called the eigenvalue-eigenvector pairs for the matrix $\boldsymbol{A}$.

The eigenvalues, λ_j, are the roots to the nth degree polynomial,

$$p_A(\lambda) \equiv \det(\lambda \boldsymbol{I} - \boldsymbol{A}). \tag{1.18}$$

If the matrix $\boldsymbol{A}$ is real-valued, and there is a complex eigenvalue, $\lambda = r + i\omega$, with an eigenvector, $\boldsymbol{V} = \boldsymbol{U} + i\boldsymbol{W}$, then there is also a complex conjugate eigenvalue $\bar{\lambda} = r - i\omega$ and a complex conjugate eigenvector $\overline{\boldsymbol{V}} = \boldsymbol{U} - i\boldsymbol{W}$, where $\boldsymbol{U}$, $\boldsymbol{W}$, r, ω are all real. This means that we can choose the constants in equation (1.16) so that the resulting solutions to the differential equation are all real. We remind the reader of the important identity:

$$e^{(r+i\omega)t} = e^{rt}(\cos \omega t + i \sin \omega t). \tag{1.19}$$

For example, referring back to our spring-mass system, the characteristic equation for the eigenvalues is

$$0 = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 + m/k. \tag{1.20}$$

Therefore the eigenvalues are purely imaginary: $\lambda_{1,2} = \pm i\omega$ where $\omega = (k/m)^{1/2}$. The mass oscillates with a frequency ω.

Now let's discuss an important aspect of linear autonomous differential equations: notions of stability. First note that $\boldsymbol{X} = 0$ is always a solution to (1.15). That is, if we start with nothing we stay with nothing. A solution that does not change with time is called an *equilibrium* point or *fixed point* of the differential equation. If we choose initial conditions close to the equilibrium point $\boldsymbol{X} = 0$, it is natural to ask if the solution will stay close and even if it will decay back to the equilibrium. Any solution to the linear differential equation can be written in terms of the general solution $\boldsymbol{X}(t)$ given by (1.16). Write $\lambda_j = r_j + i\omega_j$, where if the eigenvalue is real, $\omega_j = 0$. If each of the r_j is negative, then all solutions to the linear equation exponentially decay to zero. In this case, we say zero is *linearly asymptotically stable*. If at least one of the r_j is positive, then some solutions to equation (1.15) will grow exponentially and zero will be linearly unstable. Suppose that all the $r_j < 0$ except for one pair of imaginary eigenvalues, $\pm i\omega$. Then all components of $\boldsymbol{X}(t)$ will decay except for those corresponding to the imaginary eigenvalues, and we are left with a sinusoidally varying solution $\boldsymbol{X}(t)$ with a period of $2\pi/\omega$. We say that zero is *neutrally stable* in this case. Neutral stability is special since small changes in the coefficients of $\boldsymbol{A}$ will destroy the neutral stability and make zero either asymptotically stable or unstable.

It is useful to find conditions on the matrix $\boldsymbol{A}$ such that solutions are stable, unstable, or neutrally stable. This is a difficult problem, but it can be addressed by looking at the characteristic polynomial, $p_A(\lambda)$. Asymptotic stability is guaranteed if all the roots of the characteristic polynomial $r_j + i\omega_j$ satisfy $r_j < 0$. If we write this polynomial as

$$p_A(x) = x^n + q_{n-1}x^{n-1} + \cdots + q_1 x + q_0, \tag{1.21}$$

then it turns out that there is a simple set of conditions on the coefficients q_k that guarantee an asymptotic stability called the Routh-Hurwitz criterion. This can be found in many textbooks (e.g., Edelstein-Keshet, 2005, p. 233), so we do not state it here. Rather, we write the conditions for $n = 1, 2, 3, 4$. The state 0 is asymptotically stable if and only if

$n = 1.\, q_0 > 0.$

$n = 2.\, q_0 > 0,\ q_1 > 0.$

$n = 3.\, q_2 > 0,\ q_0 > 0,\ q_1 q_2 > q_0.$

$n = 4.\, q_3 > 0,\ q_1 > 0,\ q_0 > 0,\ q_3 q_2 q_1 > q_1^2 + q_3^3 q_0.$

In the cases, $n = 2, 3, 4$, if the last inequality is an equality (replace the $>$ by an $=$), then there is neutral stability and exactly one pair of imaginary eigenvalues, $\pm i\omega$.

It is instructive to look at the $n = 2$ case in more detail because we will need it later in this chapter. The matrix A has the form

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{1.22}$$

and the characteristic polynomial is

$$x^2 - (a + d)x + ad - bc. \tag{1.23}$$

The reader will recognize that $a + d$ is the trace (sum of the diagonals) of the matrix and $ad - bc$ is the determinant. Thus all solutions to a linear two-dimensional (2-D) autonomous system decay to zero exponentially if and only if the trace $(a + d)$ is negative and the determinant $(ad - bc)$ is positive. *These are not conditions for any higher-order system; only $n = 2$.*

In summary, linear (constant coefficient) autonomous systems have solutions that are sums of exponentials. The exponents are eigenvalues of the matrix A and their real parts determine the stability of the system about equilibrium.

1.3 Nonlinear Systems: Geometric Approach, Phase Planes

Although nonlinear systems can rarely be solved analytically in closed form, we have useful tools for characterizing their dynamic properties. These methods enable us to find parameter regimes where behaviors of interest arise, e.g., where systems change from steady-state behavior to oscillatory behavior. Without getting into specialized approximations and lengthy analytic contortions and expressions, we can gain qualitative insight from a geometric viewpoint and from dynamic systems software. The sources (Strogatz, 2001; Izhikevich, 2006; Edelstein-Keshet, 2005) are particularly oriented toward the geometric approach.

Some Definitions and Preliminaries

We can write a general nonlinear ODE in vector form:

$$\frac{d\boldsymbol{X}}{dt} = \boldsymbol{F}(\boldsymbol{X}), \tag{1.24}$$

where $\boldsymbol{X}$ is a vector of variables and $\boldsymbol{F}$ is a vector of functions that specify the rates of change for the variables; $\boldsymbol{F}(\boldsymbol{X}) = [f_1(\boldsymbol{X}), \dots, f_n(\boldsymbol{X})]$. The vector, $\boldsymbol{X}(t) = [x_1(t), \dots, x_n(t)]$ is called the state vector. For our purposes, $x_j(t)$ are real and thus the vector $\boldsymbol{X}$ lies in the n-dimensional real space, R^n. We call the space in which the variables lie the *state space* or the *phase space* of the system. For example, if there is only one variable, then the state space is the line and if there are two variables, the state space is the plane, or *phase plane*. Solution curves $\boldsymbol{X}(t)$ trace out curves in the phase space called *trajectories*. These are always one-dimensional curves parameterized by t, the time variable. One special type of trajectory is the constant one, $\boldsymbol{X}(t) = \boldsymbol{X}^*$ where $\boldsymbol{F}(\boldsymbol{X}^*) = 0$. This is the equilibrium point, steady state, or fixed point of the system. Linear systems usually have only one equilibrium point, but nonlinear systems can have many.

Another class of solutions found in neuroscience are periodic trajectories, $\boldsymbol{X}(t) = \boldsymbol{X}(t+T)$, where $T > 0$ is the period. If no other neighboring trajectories are periodic, then this trajectory is called a *limit cycle*. Limit cycles and other periodic trajectories require at least two dimensions in the phase space. Periodic trajectories (also called periodic orbits) trace out a closed curve in the phase space.

The idea of stability is a useful concept from differential equations that generalizes to nonlinear equations. A particular trajectory, $\boldsymbol{X}_p(t)$ (for example, a fixed point or limit cycle), is *stable* if any trajectories that are nearby at time t_0 remain nearby for all $t > t_0$. If any nearby trajectories diverge from $\boldsymbol{X}_p(t)$ as time increases, we say that $\boldsymbol{X}_p(t)$ is unstable. A stronger and, for neuroscience, more relevant form of stability requires in addition that as $t \to \infty$, nearby solutions converge to the orbit. For fixed points, this means that all nearby trajectories will tend to the fixed point for long times. For periodic orbits, nearby trajectories converge to the closed curve formed by the limit cycle. This type of stability is called *asymptotic stability*. The stability of equilibria is easy to determine by linearizing the equations about the equilibrium. That is, let $\boldsymbol{X}^*$ be an equilibrium solution and write $\boldsymbol{X}(t) = \boldsymbol{X}^* + \boldsymbol{Y}(t)$. Then

$$\frac{dY}{dt} = F(X^* + Y) \approx AY + R(Y) \tag{1.25}$$

from Taylor's theorem. For a one-variable system, A is the derivative dF/dx evaluated at X^*; for an nth-order system, A is an $n \times n$ matrix. $R(Y)$ is the remainder term and contains quadratic and higher-order terms. If Y is small, then the quadratic terms are much smaller than the linear terms. Since we are interested in what happens near the equilibrium point, we will assume that Y is small. [Note that $Y(t) = X(t) - X^*$ represents the deviation from the equilibrium point.] Truncating the remainder terms leaves us with

$$\frac{dY}{dt} = AY, \tag{1.26}$$

the familiar equation (1.15). This linear system (or more commonly, the matrix A) is the *linearization* of equation (1.24) around the equilibrium point X^*. The matrix A is the matrix of partial derivatives of F evaluated at the equilibrium point; the entry in row i and in column j is the partial derivative $\partial f_i(X)/\partial x_j$, the derivative of f_i with respect to x_j, treating the other variables as parameters. Now, if all of the eigenvalues of A have negative real parts, then $Y(t)$ decays to zero and there is a theorem that says that the equilibrium point, X^*, is asymptotically stable as a solution to the nonlinear system. Similarly, if any of the eigenvalues of A have positive real parts, then the equilibrium point is unstable. There is an analogous linearization and stability theory for limit cycle solutions that provides a condition for a periodic orbit to be stable.

One-Dimensional Dynamics

We now turn to a geometric description of nonlinear systems—a viewpoint that is based on trajectories and structure of the phase space. Let's demonstrate with a simple example. Consider the nonlinear first-order system

$$\frac{dx}{dt} = f(x) + I, \tag{1.27}$$

where $f(x) = x(x-a)(1-x)$, and suppose $0 < a < 1$. We start with the simple case in which $I = 0$. This nonlinear ODE has three different steady-state solutions: $x = 0, a, 1$. (This is what makes nonlinear systems much more interesting and useful; they can have many possible states.) Does this mean that the system it represents has three possible resting states? Potentially, yes; but in fact a steady state is, practically speaking, realizable only if it is stable. Applying the methods of the previous section, a steady state will be stable if the eigenvalues of the linearization have negative real parts. For a scalar model, the matrix A is a scalar, $c = f'(x^*)$, where x^* is one of the three steady states, 0, a, 1. Thus a condition for a steady state to be asymptotically stable is that $c = f'(x^*) < 0$. For our example, $f'(x) = -3x^2 + 2(a+1)x - a$, so that $f'(0) = -a < 0$ and $f'(1) = a - 1 < 0$ since a is between 0 and 1. Thus both $x = 0$ and $x = 1$ represent stable steady states. For the case $x^* = a$ we find that this steady state is unstable since $f'(a) = a(1-a) > 0$; thus, we would not physically see this state. This simple system is bistable. Starting from an initial condition near $x = 0$ or $x = 1$, the system approaches one of the two stable states. If the parameter I is nonzero, then finding the equilibrium points is much harder (we have to solve a cubic equation) and it provides little insight. For this reason, we turn to the geometric approach.

The x-axis is the phase space for this one-variable dynamic system. Think of $x(t)$ as a point moving, as t increases, along the x-axis. Next, plot the rate function $f(x)$ on the vertical axis. Immediately, we know the direction in which $x(t)$ moves, and we indicate that flow with arrows on the x-axis (figure 1.1A). Where $f(x) > 0$, the arrow at x points rightward (since dx/dt is positive) and where $f(x) < 0$, the arrow is leftward. If $x(0)$ lies to the left of position $x = a$, then $x(t)$ moves smoothly toward $x = 0$, asymptotically approaching this stable steady state as $t \to \infty$. On the other hand, if $x(0)$ lies to the right of $x = a$, then $x(t)$ drifts toward $x = 1$. The two steady states $x^* = 0, 1$ are *attractors* and the phase space is divided into their two domains of attraction by the unstable steady state $x^* = a$. The bistable system shows a threshold behavior (figure 1.1B). It is rather elegant that the geometric, phase-space treatment gives us an immediate global view.

The preceding analytic approach was really only a local description, based on linearizing $f(x)$ around x^*. Although limited, it did give us the local rate of (exponential) approach or repulsion. The geometric approach shines brightest when we introduce the parameter I. As noted earlier, it is tedious to find the equilibrium points and their stability analytically for the cubic equation (1.27). However, with the geometric method, the addition of I does not change the methods and ideas. If $I > 0$, then the cubic picture in figure 1.1A is raised (see figure 1.1C) and if $I < 0$, it is lowered. Two interesting things happen as I is changed. As I increases, the two equilibria corresponding to 0, a merge when $I = I_2$ and no longer exist for $I > I_2$. The model has only one stable equilibrium point. Since the behavior is qualitatively different (monostable versus bistable), we say that a *bifurcation* takes place at $I = I_2$. If I is near I_2, two distinct behaviors are observed, depending on whether I is above or below I_2. As I is lowered, a bifurcation occurs at I_1 where the equilibria corresponding to a, 1 merge and disappear. We summarize the behavior in the compact diagram shown in figure 1.1D, which is a bifurcation diagram. This diagram shows only the steady states and indicates if they are stable (solid) or unstable (dashed). For a fixed value of I between I_1 and I_2, there are three steady states, while for $I > I_2$ or $I < I_1$ there is only one.

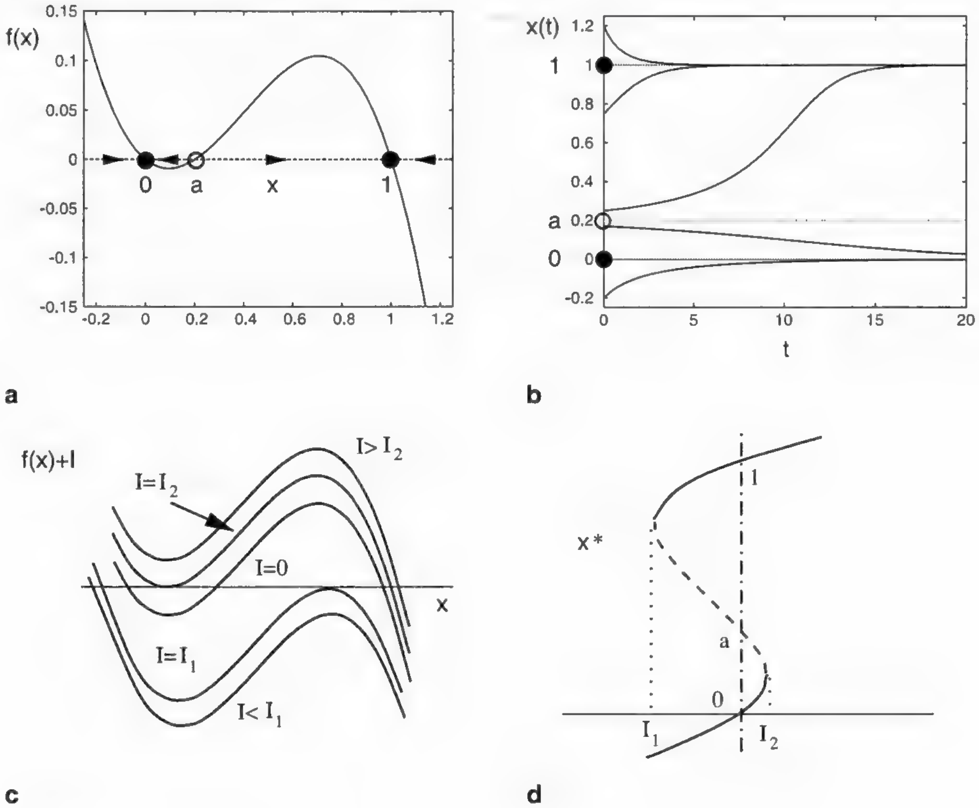


Figure 1.1
One-dimensional phase space.

Multidimensional Dynamics; Phase Planes

Clearly, the one-dimensional phase line is good only for one-variable systems. For higher-dimensional systems, the geometric methods are powerful but difficult to generalize except when the system is planar. Surprisingly, much of the behavior of neurons can be captured by looking at two-dimensional dynamics. So, let us jump into two dimensions and introduce phase-plane analysis. Consider the general two-variable system before we get specific:

$$\begin{aligned}\frac{dx}{dt} &= f(x, y)\\ \frac{dy}{dt} &= g(x, y).\end{aligned} \tag{1.28}$$

We interpret the pair of rates (f, g) geometrically. They specify the local velocity vector, the direction of motion at position (x, y) in the plane. At points where $f(x, y) = 0$, the x-direction velocity is zero. The points for which $f(x, y) = 0$ define a curve implicitly, a special curve in the plane, the x-nullcline. When a trajectory crosses the x-nullcline, the flow is vertical; $x(t)$ would be a local maximum or minimum. Similarly, the points where $g(x, y) = 0$ define the y-nullcline, along which the flow is horizontal. By drawing these two curves and then looking at a few local velocity vectors, we can learn a lot about the qualitative behavior of the 2-D dynamic system. Furthermore, at (x, y) points where the two nullclines intersect, both dx/dt and dy/dt vanish, so that these are steady states for the two-dimensional system. As with the one-dimensional system, we can immediately see the equilibrium points in the system (1.28) by graphing the nullclines and observing intersections. Unlike the one-dimensional system, the stability of the steady states cannot in general be determined by inspection. With interactive dynamic systems software packages such as XPPAUT (see the software appendix at the end of this volume) the flow field and nullclines appear at the push of a button.

Suppose that we have found a point $(x, y) = (x^*, y^*)$ that is an equilibrium or steady state: $f(x^*, y^*) = g(x^*, y^*) = 0$ (easily viewed in the phase plane by looking at the intersections of the nullclines.) Unlike the one-dimensional example, it is hard to discern stability from the phase plane alone so that we are forced to evaluate stability using the linearization techniques. We define

$$A = \begin{pmatrix} \partial f(x^*, y^*)/\partial x & \partial f(x^*, y^*)/\partial y \\ \partial g(x^*, y^*)/\partial x & \partial g(x^*, y^*)/\partial y \end{pmatrix} \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{1.29}$$

to be the linearized system around an equilibrium point. From section (1.1) we see that the equilibrium will be stable if and only if $a + d < 0$ and $ad - bc > 0$.

As an example, consider the following model:

$$c\frac{dV}{dt} = I - g_{\mathrm{L}}(V - E_{\mathrm{L}}) - g_{\mathrm{Ca}}m_\infty(V)(V - E_{\mathrm{Ca}}) - g_{\mathrm{K}}w(V - E_{\mathrm{K}}) \equiv cf(V, w) \tag{1.30}$$

$$\frac{dw}{dt} = [w_\infty(V) - w]/\tau_w(V) \equiv g(V, w), \tag{1.31}$$

which is a one-compartment model for the membrane potential of a barnacle muscle. [This so-called Morris-Lecar model (Morris and Lecar, 1981) is taken through its paces in Rinzel and Ermentrout (1998).] Here the x-axis will be the membrane potential, V, and the y-axis will be the fraction w of potassium channels that are open. (Note that the notation $\equiv$ indicates that we are defining a symbol for an expression.)

In figure 1.2A we illustrate the geometric methods by showing the phase-plane portrait of equations (1.30)–(1.31); it is representative for models of many excitable

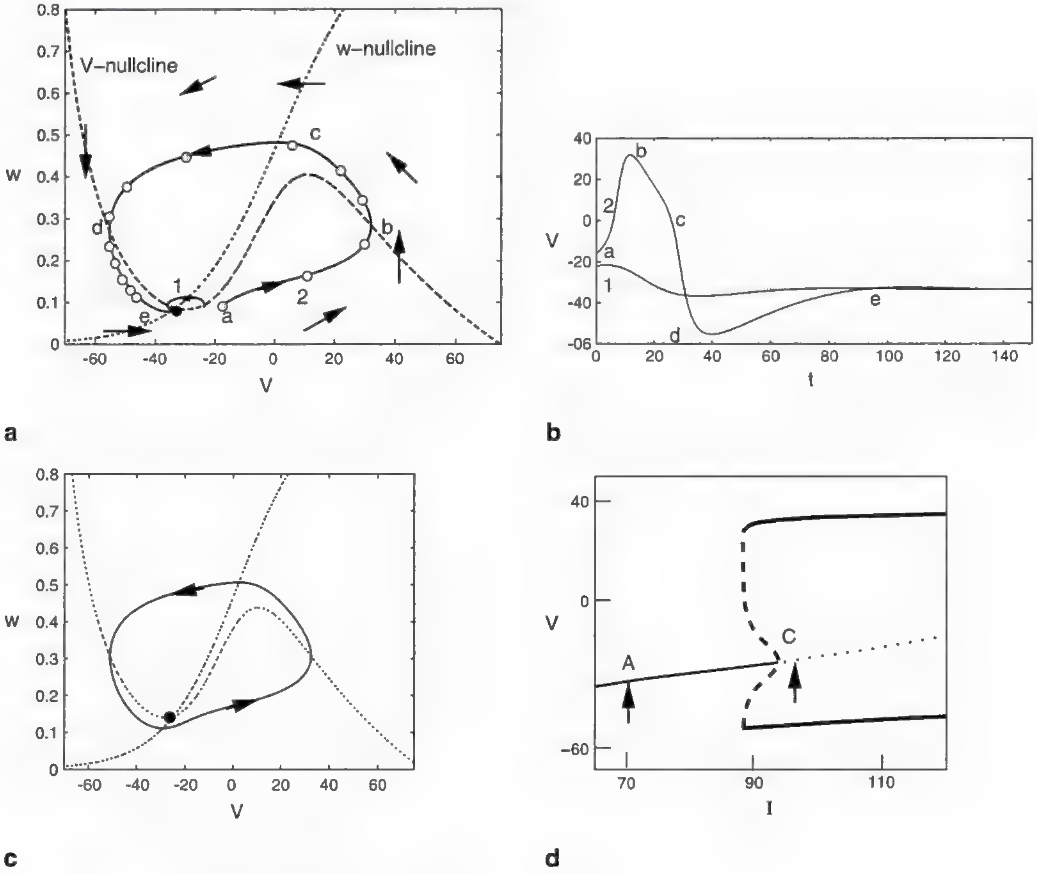


Figure 1.2
Two-dimensional phase space and excitability.

systems, such as two-variable models for an action potential and membrane electrical excitability. The point (V^*, w^*) where the two nullclines intersect corresponds to a steady state or fixed point of the system (shown as a filled circle). There, both dV/dt and dw/dt are zero. In this example, (V^*, w^*) is unique and is stable. We might refer to it as the resting state of the system.

Now let's look at the flow. Each point in the phase plane is associated with a direction found by evaluating $dV/dt = f(V, w)$ and $dw/dt = g(V, w)$ at the given point. If both dV/dt and dw/dt are positive at a point (V, w), then they will be increasing at this point and we draw an arrow pointing up and to the right. In this model, dV/dt is positive for points below the V-nullcline and dw/dt is positive below the w-nullcline. Hence, the nullclines partition the phase plane into regions where the flow is in one of four directions: up and to the right ($dV/dt > 0, dw/dt > 0$), up and to the left ($dV/dt < 0, dw/dt > 0$), down and to the right ($dV/dt > 0, dw/dt < 0$), and down and to the left ($dV/dt < 0, dw/dt < 0$). For points on the nullclines, either $dV/dt = 0$ or $dw/dt = 0$ so the arrows are respectively vertical or horizontal.

Figure 1.2A shows some arrows indicating the directions at selected points as well as to trajectories (solutions) starting at different points in the plane. The velocity arrows show a generally counterclockwise flow. Starting from any initial position, the trajectory will eventually return the system back to the resting state. This is a powerful statement: the resting state (x^*, y^*) is a global attractor for the system. Moreover, from the layout of these nullclines and a few flow arrows, we see the essential features of excitability. A small perturbation from resting (say, in V) produces a response that does not grow in V; the trajectory returns directly to rest (the trajectory labeled 1). If a large enough initial V-perturbation pushes the phase point rightward so that (at $t = 0^+$) it lies below the V-nullcline's hump, then $dV/dt > 0$ and we get amplification and an action potential.

Figure 1.2B shows the voltage as a function of time for the subthreshold (1) and suprathreshold (2) initial conditions. For the suprathreshold initial conditions generating a spike, we label the parts of the action potential in figure 1.2B and their corresponding points in the phase plane in figure 1.2A. The suprathreshold trajectory takes off, heading toward the right branch of the V-nullcline. This demonstrates the thresholdlike behavior and the spike's upstroke and occurs between points a and b. The trajectory turns vertical when it reaches the V-nullcline, corresponding to the peak voltage of the spike (at point b). Next, the flow carries the phase point to the upper left (point c), with w increasing and leftward drift constrained by the V-nullcline. Note that the trajectory cannot cross (leftward) back over the V-nullcline because of the rightward flow over there. This is the phase of growing negative feedback w (potassium current activation) during a spike's depolarized (plateau) phase (between b and c). However, after the trajectory rises to above the V-nullcline's upper "knee," then it is free to dramatically turn leftward (dV/dt is strongly negative there). The spike's downstroke proceeds, and maximum w occurs during the downstroke when the trajectory crosses the w-nullcline. With the fall of V to its minimum (point d), we return to rest during this recovery phase (point e).

The phase-plane plots do not provide information about the time it takes to go from one point to another: they are purely geometric concepts. However, one can plot a few gray circles as isolated points along the trajectory to give a sense of the timing between points. In figure 1.2A we add a few circles to illustrate the position every 5 ms for the first several points. During the upstroke and downstroke, the distance between successive points is larger than along the plateau and recovery. This type of behavior is typical in neural models where the voltage dynamics are typically much faster than the recovery.

The geometric view has enabled us to approximately predict the spike's trajectory without actually computing it. Many other properties of the system can be deduced from the phase-plane analysis, such as absolute and relative refractoriness, postinhibitory rebound, and repetitive firing (e.g., see Izhikevich, 2006, ch. 4). For example, the effect of a steady (depolarizing) stimulus current is to translate the V-nullcline upward (with a little distortion). As the stimulus increases, the equilibrium point moves toward the right (more depolarized) and upward (more recovery) along the w-nullcline (always, by definition, at the nullclines' intersection point). For an adequate stimulus, (V^*, w^*) will be on the middle branch of the V-nullcline (figure 1.2C). If w is relatively slow, then we can immediately conclude from the phase plane that (V^*, w^*) is unstable in this situation. The flow in the V-direction is strongly away from (x^*, y^*); it points leftward just to the left of (x^*, y^*) and rightward just to its right. In this case, trajectories converge onto the closed orbit of repetitive spiking; here, this *limit cycle* is the global attractor. Periodic solutions to the differential equation are represented by a closed trajectory in phase space. All solutions to the ODE (except for those starting at the unstable equilibrium point shown in the figure as a filled black dot) will converge to this closed trajectory.

Unlike the changes seen in figure 1.1, here the number of equilibria does not change as the parameter varies, but the stability of the equilibrium point *does* change. The change in stability of (V^*, w^*) can be characterized analytically. One linearizes about (V^*, w^*) and finds that for a critical stimulus level, the local behavior changes from a damped to a growing oscillation. The eigenvalues are a complex pair, $r \pm i\omega$, for which r changes from negative to positive at criticality. When an equilibrium point changes from stable to unstable via a pair of complex eigenvalues, then we expect to find limit cycles for values of the parameter near the transition point. This fact is a consequence of the *Hopf bifurcation theorem* (Edelstein-Keshet, 2005, p. 341).

Figure 1.2D shows the bifurcation diagram for equations (1.30–1.31) as the current I increases. The maximum and minimum values of the voltage are shown along the y-axis while the current is shown on the x-axis. Figure 1.2A corresponds to a current of 70 $\mu A/cm^2$ (shown in figure 1.2C by the arrow labeled A). The thin solid line represents the voltage component of the stable equilibrium point as the current varies. The thin dashed line represents the unstable equilibrium. At about $I = 95$ $\mu A/cm^2$, the equilibrium becomes unstable. At this point, a new solution appears; it is an unstable limit cycle (thick dashed line). The maximum and minimum values of the voltage are plotted for each current level. This unstable limit cycle exists only for a small range of currents between around 87 and 95 $\mu A/cm^2$. For currents beyond 95 $\mu A/cm^2$, there is a unique stable limit cycle represented by the thick solid curve in figure 1.2D. For currents between 87 and 95 $\mu A/cm^2$, there is both a stable equilibrium and a stable limit cycle.

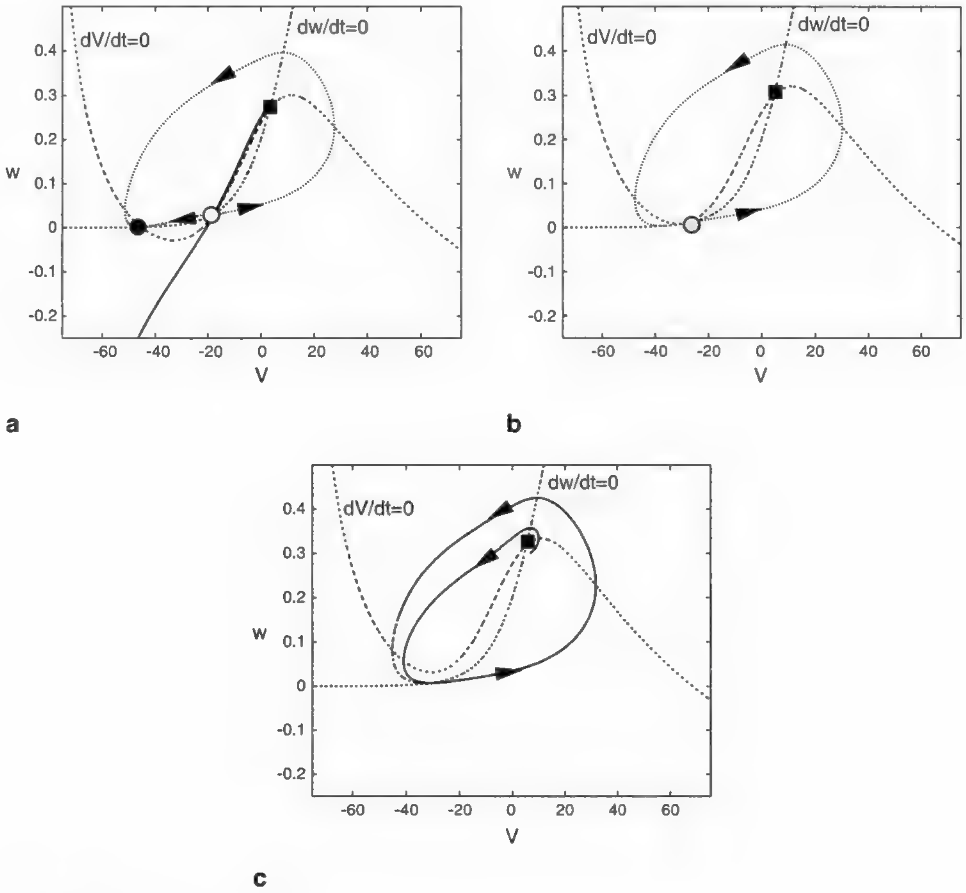


Figure 1.3
Saddle-node excitability.

Amazingly enough, there are only a very few ways in which an equilibrium point can give rise to a limit cycle as a parameter varies. The Hopf bifurcation described here is one such example, which occurs commonly in models of neurons. In fact, the Hopf bifurcation is the mechanism by which the Hodgkin-Huxley squid axon model makes the transition from being at rest to repetitive firing. There is another common mechanism through which neurons go from a resting state to regular firing called the *saddle node on an invariant circle* or simply, SNIC bifurcation. Figure 1.3 illustrates this bifurcation in the model (equations 1.30 and 1.31) with different parameters from figure 1.2. At low values of the current in figure 1.3A there are three equilibrium points: a stable node (the black circle), an unstable node (black square), and a saddle point (white circle). Contrast this with figure 1.2, where there is always one equilibrium point. The resting state is the stable node. A diagonal line enters the saddle point and forms a true threshold curve for the model. Any perturbations that move the voltage to the right of this curve from the resting state will result in an action potential. As the current increases, the saddle and the stable node coalesce at a saddle-node bifurcation (figure 1.3B, gray circle). There is now a limit cycle, but it contains an equilibrium point, so that it has an "infinite period" or zero frequency. A small increase in the current lifts the V-nullcline past the w-nullcline (figure 1.3C), allowing the system to fire repetitively. Since lowering the current can bring the nullclines arbitrarily close together, it can take a very long time to get through the narrow gap and thus the frequency of this model can be arbitrarily low. No bistability is possible near the SNIC, in contrast to the Hopf bifurcation. Many models for regularly spiking pyramidal cells have this bifurcation. Further coverage on the SNIC, Hopf, and other bifurcations in neuronal models can be found in Izhikevich (2006) and Rinzel and Ermentrout (1998).

We have only scratched the surface of dynamic possibilities, even for 2-D ODE systems. Our main goal here has been to demonstrate the benefits and insights obtained from a geometric approach and how you can predict qualitative behavior if you develop a little intuition about how the nullclines depend on system parameters. For example, you could easily imagine that changing some parameters could distort the nullclines and lead to multiple intersections (therefore, multiple steady states) and possibilities for multistability and SNIC bifurcations.

Although the phase-plane treatment is completely thorough only for 2-D systems, one can use the concepts to gain insight into the behavior of some higher-order systems; e.g., by looking at projections that show a fast and a slow variable. Also, if one variable z is very slow, then one may consider the phase plane as parameterized by z, and snapshots or animations of the phase plane can be illuminating. The bifurcation phenomena (Hopf, SNIC, saddle node) are general, occurring in higher-dimensional systems, beyond 1-D and 2-D cases.

1.4 Space, Time, and Partial Differential Equations

One issue confronting a modeler is when to use ODEs for a system. We consider two situations where there might be some questions. Suppose that we wish to model the electrical activity of a neuron with dendrites or an axon.

A typical neural model involves many variables, such as the potential, the calcium concentration, and the gating variables. Is it reasonable to assume that the neuron is equipotential? That is, do we expect the potential in, say, the soma, to be the same as that halfway along the apical dendrite? If the neuron is electrically compact, then this is probably not a bad approximation. However, if channel densities vary wildly or the neuron is spatially extended, then it is probably better to subdivide the neuron into compartments and write differential equations for each of these (see chapter 11, section 1). Reducing a neuron to a single compartment is making the implicit assumption that the neuron is space clamped. Similarly, if we are modeling chemical reactions in a beaker, a common assumption is that the chemical species are well mixed, so that there are no spatial differences. Then it is possible to write a single differential equation for each of the reacting species (see chapter 3). In both the neuron example and the chemical example, we can write a single differential equation for each variable if we believe that the spatial differences are not important. In many modeling studies, this is a good first step. On the other hand, if we think that spatial differences matter, then it is necessary to write equations for the variables at many (possibly infinitely many) spatial locations. Figure 1.2A showed an example of a typical excitable membrane model.

Suppose we want to model a piece of axon to assess how long it takes the action potential to propagate over some small distance. Then, clearly, we should not treat the axon as space clamped or well mixed. Instead, we subdivide the axon into N segments or compartments of length Δx and assume that they are connected to their nearest neighbors via the intracellular resistance. Each compartment obeys equations (1.30)–(1.31). We can thus write:

$$\begin{aligned} c\frac{dV_j}{dt} &= \frac{D}{\Delta x^2}(V_{j+1} - 2V_j + V_{j-1}) + I_0 - I_{\text{ion}}(V_j, w_j) \\ \frac{dw_j}{dt} &= \frac{w_\infty(V_j) - w_j}{\tau_w(V_j)}, \end{aligned} \tag{1.32}$$

where I_0 is the applied current (applied to every compartment) and

$$I_{\text{ion}}(V, w) = g_L(V - E_L) + g_{\text{Ca}} m_\infty(V)(V - E_{\text{Ca}}) + g_{\text{K}} w(V - E_{\text{K}}) \tag{1.33}$$

represents the membrane ionic currents in the axon. We mention a few interesting points that these equations bring out; for other details, see chapter 10, section 1. First, $D = a/(2 \times R_{\text{axial}})$ where a is the radius of the dendrite and R_{axial} is the specific axial resistivity (typically measured in Ω-cm, so that D has dimensions of conductance). Second, we have written this equation in terms of an index j corresponding to the jth compartment. Interactions occur with the neighbors, $j \pm 1$. So, it is natural to ask what happens at $j = 1$ and at $j = N$. For example, at $j = 1$, there is a reference to V_0 so that in order to make sense of the equation when $j = 1$, we have to specify V_0. Similarly, at $j = N$, we need to specify V_{N+1}. The values for these "boundary" points are the boundary conditions for the model. There are many choices for boundary conditions possible.

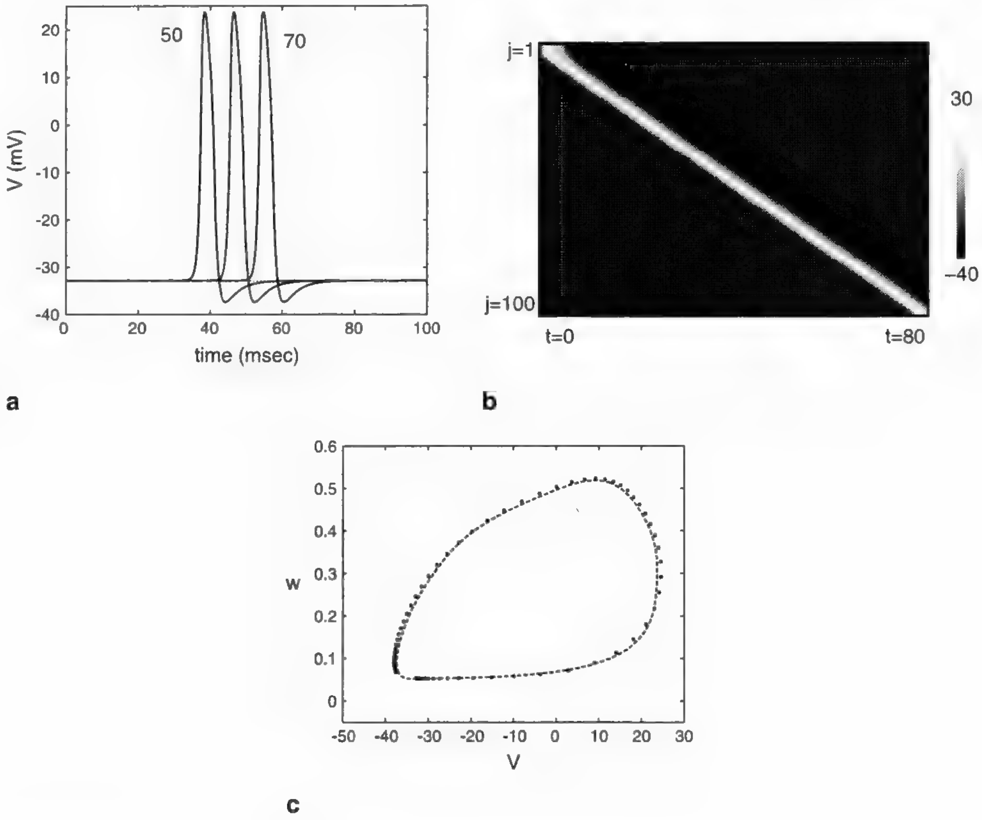


Figure 1.4
Partial differential equations.

Figure 1.4 shows a simulation of an axon broken into 100 compartments, $D/\Delta x^2 = 1$. We have set $V_0 = V_1$ and $V_{101} = V_{100}$ for this simulation, which corresponds to clamping the current flow out the ends to zero. We start all compartments at their resting value except the first three, which start above the threshold. The result is a wave propagating to the right. Unlike the space-clamped equation, activity persists for a much longer time (80 ms compared with 10 ms) and if the cable were longer, the activity would persist for an even longer time. Figure 1.4A shows the potentials in three different compartments, 50, 60, and 70. The time difference between the action potential at $j = 50$ and $j = 60$ is exactly the same as that between $j = 60$ and $j = 70$. This is the hallmark of a traveling wave.

Figure 1.4B shows the full space-time picture. Time increases from left to right and compartment number increases from top to bottom. The potential is shown as a grayscale and traces out a straight line except near the ends of the cable. If we fix the length of the axon to be l and let the number of compartments, N, go to infinity (that is let $\Delta x = l/N$ go to zero), then equation (1.32) becomes a *partial differential equation (PDE)*:

$$
\begin{aligned}
c\frac{\partial V(x,t)}{\partial t} &= D\frac{\partial^2 V(x,t)}{\partial x^2} + I_0 - I_{\text{ion}}[V(x,t), w(x,t)] \\
\frac{\partial w(x,t)}{\partial t} &= \frac{w_\infty[V(x,t)] - w(x,t)}{\tau_w[V(x,t)]}.
\end{aligned}
\tag{1.34}
$$

As with an ordinary differential equation, the equations define the rates of change of the voltage and the potassium channels at each spatial location $0 < x < l$. The interaction in space is via the term $\partial^2 V/\partial x^2$, which represents two derivatives with respect to space. As with equation (1.32), we need to define conditions at the ends of the axon, $x = 0$ and $x = l$. The boundary conditions corresponding to the simulation in figure 1.4 can be written as

$$
\frac{\partial V(0,t)}{\partial x} = \frac{\partial V(l,t)}{\partial x} = 0. \tag{1.35}
$$

Physically, this means that no current can flow out the ends of the cable. In addition to the boundary conditions, initial conditions are also required. That is, $V(x,t_0) = V_0(x)$, $w(x,t_0) = w_0(x)$, where V_0, w_0 are some functions. As with the ODE, under reasonable assumptions there is a unique solution to this PDE for some positive time.

Looking at figure 1.4A, it would appear that the potential at a point (x_1, t_1) is exactly the same as the potential at a point (x_2, t_2) where $v \equiv (x_2 - x_1)/(t_2 - t_1)$ is constant (at least away from $x = 0$ and $x = L$). That is, the action potential takes the form of a constant-velocity wave moving to the right. The velocity is the constant v, which for the present simulation is roughly 1.25 space units per millisecond (that is it takes 80 ms to travel 100 compartments). This observation allows us to convert the PDE into a special kind of ODE. First, we make the axon infinite in length since we want the wave to travel forever. This is a mathematical conceit that allows us to simplify the equations a great deal. The way to think of this is that when the traveling action potential is in the middle of the axon, the behavior at the ends is essentially irrelevant to the dynamics in the middle. Let $y = x - vt$ be a moving coordinate system; a large and negative y corresponds to regions in space and time where the wave has already passed (points at the left of the line in figure 1.4B) while a large and positive y corresponds to regions that have not yet experienced the activity. Let $V(x,t) = u(y)$ and $w(x,t) = z(y)$ be the potential and the potassium activation in the moving frame. Then the PDE, equation (1.34), can be written as an ODE:

$$
\begin{aligned}
-vc\frac{du}{dy} &= D\frac{d^2u}{dy^2} + I_{\text{ion}}(u, z) \\
-v\frac{dz}{dy} &= \frac{w_\infty(u) - z}{\tau_w(u)}.
\end{aligned}
\tag{1.36}
$$

Unlike the ODEs we have encountered before, we do not want to specify initial conditions for this. Instead we want the voltage and the potassium gates to be at rest at the far ends of the axon and we want to specify a particular value of the potential in the middle of the cable, say, $V = 0$. This latter condition simply sets the origin for the wave since it is translation invariant. Thus, we want

$$u(\pm\infty) = V_{\text{rest}}$$

$$z(\pm\infty) = w_{\text{rest}}$$

$$u(0) = 0.$$

As we noted earlier, it is always possible to satisfy the last condition as long as $u(y)$, $z(y)$ is not identical to the resting state. ODEs where conditions are specified at different values of the independent variable are called *boundary-value problems* (BVP) and are considerably harder to solve both analytically and numerically. Equation (1.36) can be written as three first-order equations (as we did in the earlier part of this chapter). In general, three ODEs means we can specify at most three conditions. However, we have four conditions (both u, z must go to rest as $y \to \pm\infty$). However, there is a free parameter, the velocity, v. Thus we want to know if there is a value of v such that the solutions to equation (1.36) satisfy the four boundary conditions.

Unlike initial-value problems, there is no guarantee that there is a solution to the BVP. Through a method called shooting, it is possible to estimate the value of the velocity $v \approx 1.34$ so that the four conditions at $\pm\infty$ are met. We remark that this value is within 10% of our crude estimate from figure 1.3A. Figure 1.3C shows the projection onto the $V - w$ phase plane of the solution to the compartmental model at $j = 50$. Superimposed on this as a series of dots is the solution to equation (1.36). Thus we see that while shooting requires some work, it provides a good description of the discretized solutions to the full PDE. In their pivotal paper on the squid action potential, Hodgkin and Huxley (1952) also used shooting to estimate the velocity of the action potential since solving the full PDE was beyond the computational capabilities at the time. In conclusion, PDEs are necessary when spatial differences matter; PDEs require both initial and boundary conditions; and certain forms of solutions to PDEs can be reduced to ODEs.

1.5 Noise

Let us turn to chemicals reacting inside a cell (see chapter 3) or channels opening and closing in a small patch of membrane (see chapter 5). Within the cell, some chemical species may be represented by as few as one or two molecules, such as messenger-RNA (mRNA) for some protein. In this case, it does not make any sense to write a differential equation for the concentration of the mRNA. Similarly, in a tiny patch of membrane (e.g., on the head of synaptic spine), there may be only a few dozen channels and so it becomes questionable whether to treat the fraction of open channels as a continuous variable. Rather, we may have to treat each species or channel as a random individual. (The randomness arises because the opening and closing of individual channels is typically a stochastic process.) For a large patch of membrane or in a larger environment, there are many more channels or molecules and thus, if they are acting independently, it makes sense to treat the fraction or concentrations as continuous variables and thus differential equations. However, in low concentration or when there are few channels, the fluctuations [which die off like $1/(M)^{1/2}$, where M is the number of molecules or species] can dominate the behavior and e.g., produce finite-time extinction or oscillations where none occurred for the continuous model. Still, we may think of this as a well-mixed system with only a finite number of channels or molecules.

Returning to our old friend the Morris-Lecar equation, we replace the dynamics for the potassium channel activation in equation (1.31) with a stochastic value determined by the following dynamics: we suppose that there are M independent potassium channels. Typically, channel opening and closing is treated as a Poisson process with a rate that is a function of the voltage. This means that if the rate of opening is α, then the probability that a channel will go from open to closed in the time interval Δt is $\alpha \Delta t$ for small-enough time intervals. Similarly, an open channel will return to a closed state with the probability $\beta \Delta t$. If α or β is large, then we must make the time increment Δt very small. We use parameters exactly as in figure 1.2A, but have only eighty potassium channels that are governed by the random opening and closing of the gates with

$$\alpha(V) = w_{\infty}(V)/\tau_w(V)$$

$$\beta(V) = (1 - w_{\infty}(V))/\tau_w(V).$$

At any given time there are, say, p open channels so that $w = p/M$ represents the fraction of open channels and is thus used for the w in the voltage equation (1.30). Figure 1.5A shows a typical simulation of the stochastic model. Note that the fluctuations are sufficiently large to cause the neuron to fire spontaneously even though there are no stimuli. Zooming into one of the spikes (figure 1.5B), we can see the dis-

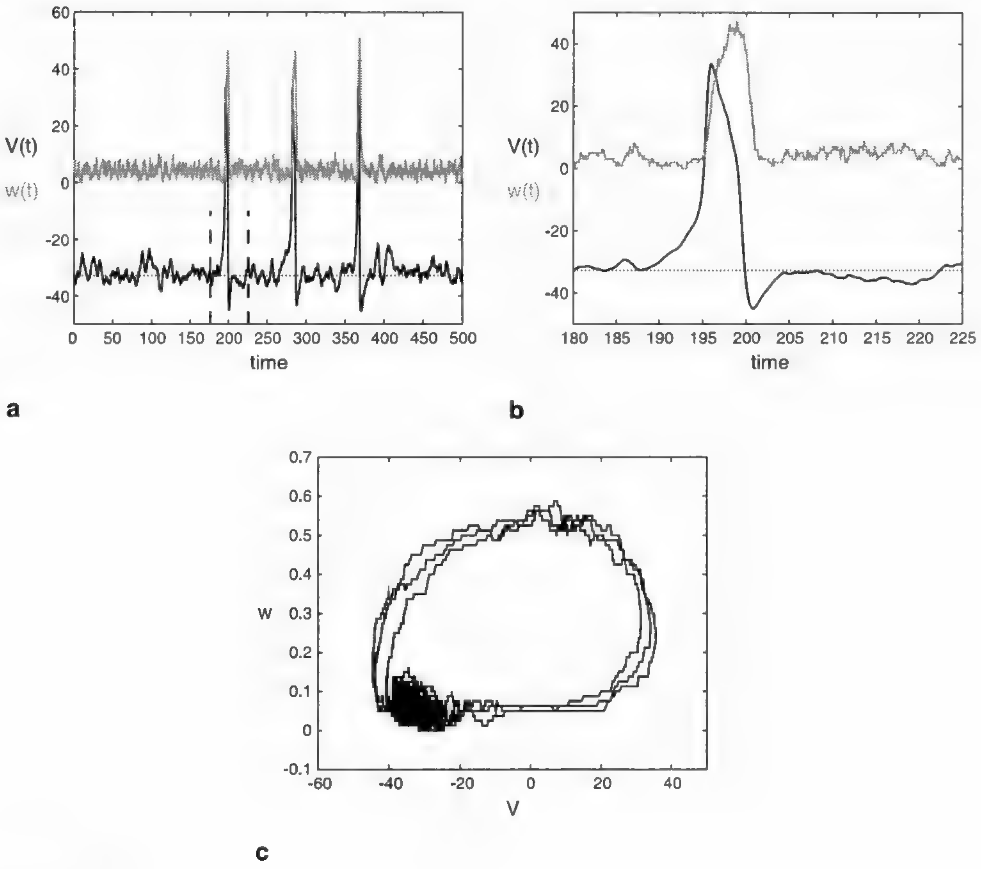


Figure 1.5
Noisy ODEs.

crete increments taken by $w = p/M$. This is even more apparent in the phase-plane plot shown in figure 1.5C; w is constant while V moves according to the deterministic dynamics of equation (1.30). With fewer channels, say twenty, the membrane is much more active and instead of producing spikes at a rate of about 6 Hz, it produces them at a rate of 30 Hz. Thus, in some instances, stochastic fluctuations matter a great deal in the qualitative behavior of the modeling. Here, with deterministic dynamics, the neuron is "dead," but with small numbers of stochastic channels, it springs to life (similar behavior is seen in the Hodgkin-Huxley model; Strassberg and DeFelice, 1993).

1.6 Numerical Methods

Very few differential equation models for neural behavior admit a closed-form solution in terms of simple functions. Thus most neural models are simulated by solving them numerically on a computer. The numerical solution of ODEs is a mature field and there are many good packages available for solving initial-value problems. Boundary-value problems are more difficult to solve and their solution remains an area of active research. PDEs are even more difficult to solve. For the kinds of PDEs in neuroscience, such as models for dendrites and axons, the easiest way to solve them is to divide them into small isopotential compartments and then solve the resulting large system of ODEs (see chapter 11, section 1). This is precisely what programs like NEURON and GENESIS do to simulate reconstructed neurons (see the software appendix).

ODEs

Here we briefly discuss methods for solving ODEs using a one-dimensional model as an example:

$$u' = G(u, t). \tag{1.37}$$

The easiest way to approximate this is to recall the finite-difference approximation:

$$\frac{\Delta u}{\Delta t} = G(u, t), \tag{1.38}$$

which says that the change in u is approximately equal to

$$\Delta u \approx \Delta t G(u(t), t). \tag{1.39}$$

Let us (for notational convenience) denote Δt by the parameter h, which we will call the step size of the approximation. Suppose we know u at t_0. Let $t_1 = t_0 + h$ and more generally, $t_n = t_0 + nh$. Then we can approximate the function $u(t)$ at these values, $u_0 = u(t_0), u_1 = u(t_1), \dots, u_n = u(t_n)$, by the iteration:

$$u_{n+1} = u_n + hG(u_n, t_n). \tag{1.40}$$

This iterative scheme, called *Euler's method*, is the easiest to implement and is therefore quite popular although it is one of the most inefficient methods; in order to be accurate, h must be very small. At each step, the approximation gives an error that is roughly of size h^2, so that at the end of a simulation the error is roughly proportional to h.

An improvement on Euler's method requires making multiple evaluations of the right-hand side, $G(u, t)$, for each step taken. For example, the well-known *Runge-*

Kutta method requires four evaluations at each time step. (The Runge-Kutta method is essentially Simpson's rule for integration; another simple method requiring two right-hand-side evaluations is *Heun's method* and is equivalent to the trapezoidal rule.) The reader might well wonder why one would want to do more work. It turns out that the Runge-Kutta method gives an error that is proportional to h^5 at each step and proportional to h^4 at the end of the simulation. Suppose that you wanted an error of 10^{-4} at the end of a simulation that lasted, say, for 1 s. So for Runge-Kutta, you take $h = 0.1$ and it requires 40 evaluations of the right-hand side. Euler requires $h = 0.0001$ for the same accuracy and thus needs 10,000 evaluations of the right-hand side. So, for a little extra coding work, you get an incredible increase in speed.

Since Runge-Kutta requires four right-hand-side evaluations, it is called a *fourth-order method.* By comparing the results of, say, a fourth-order method and a fifth-order method, it is possible to get an excellent estimate of the error for that time step. If the user prescribes a fixed maximal error, it is then possible to change the time step during the computation to stay within the error. When the equations are rapidly changing (G is large), take small steps, but when G is small, take large steps. Writing so-called adaptive step-size solvers is not as easy as the fixed-step integrators, but they can dramatically increase the speed of solving ODEs. There are many adaptive step-size packages available, so most users need not worry about this.

Accuracy is one important issue in solving ODEs. There is another important issue, *stability*, which roughly describes whether the iteration scheme will converge to a solution near the true solution. Consider

$$x' = -ax, \quad x(0) = 1, \tag{1.41}$$

where $a > 0$ is large. The solution to this is $x = e^{-at}$. The Euler approximation to this is

$$x_{n+1} = x_n - ahx_n = (1 - ah)x_n. \tag{1.42}$$

Suppose that $ah > 1$ so that $1 - ah < 0$. This means that at each step the sign of x_n switches. This is far removed from the true behavior. One way to rectify this is to make h very very small, but this will then take a very long time to simulate. Equations where there is an equivalent of a very large value of a (e.g., some components in a system relax much faster than others) are called stiff; they call for different methods. Suppose that we change the Euler method slightly to

$$x_{n+1} = x_n - ahx_{n+1}. \tag{1.43}$$

Instead of putting x_n on the right-hand side, we put x_{n+1}. Solving this for x_{n+1} yields

$$x_{n+1} = \frac{x_n}{1 + ah}. \tag{1.44}$$

This sequence of iterations stays positive and decays to zero no matter how large a is for any step size h; this method is unconditionally stable. Applying this to our general first-order equation, we must solve

$$u_{n+1} = u_n + hG(u_{n+1}, t_{n+1}). \tag{1.45}$$

We immediately see a problem: we have to solve a possibly nonlinear equation for u_{n+1}:

$$u_{n+1} - hG(u_{n+1}, t_{n+1}) = u_n. \tag{1.46}$$

Solving nonlinear equations is a major area of research in numerical analysis, so that this method, called backward Euler, is best left to the professional. Typically, if h is not too large, we have a good guess for u_{n+1}, namely, the previous value, u_n, so that one can use Newton's method to solve for u_{n+1}. Newton's method requires that we also know the derivative of the function G with respect to u, but this can be done numerically. In general for N-dimensional systems, we have to solve a large system of nonlinear equations. Methods that put x_{n+1} into the right-hand sides are called implicit or backward methods, in contrast to explicit or forward methods. The latter are much easier to code but can fail miserably with certain classes of ODEs that commonly arise. Gear methods and their descendants, such as LSODE and CVODE, combine adaptative and implicit methods to solve stiff ODEs quickly and accurately. More description of numerical methods for simulating differential equations can be found in Mascagni and Sherman (1998).

Boundary-Value Problems

Boundary-value problems arise in cases when we are looking for steady states of PDEs or, as earlier, when looking for traveling wave solutions to nerve axon equations. As we have noted, unlike initial-value problems, there is no guarantee of the existence of solutions. So before embarking on numerical solutions of BVPs, one should first be sure that there is a solution! There are two general strategies for solving BVPs: shooting and collocation. We illustrate the two ideas using a simple steady-state cable equation:

$$-V(x) + \lambda^2 \frac{d^2V}{dx^2} = 0 \quad 0 < x < 1, \tag{1.47}$$

with

$$V(0) = 1, \quad \frac{dV}{dx}(1) = 0. \tag{1.48}$$

Here λ is not the symbol for an eigenvalue, but is the cable's space constant: $\lambda = [aR_m/(2R_{\text{axial}})]^{1/2}$; R_m is the specific membrane resistivity (in units of Ω-cm^2) (see chapter 10, section 1).

The idea of shooting is as follows: This is a second-order ODE so we need to specify an initial condition and the derivative at $x = 0$. Choose $V(0) = 1$ to satisfy the initial condition at the left. Now, using the notation V' for differentiation with respect to x, we choose $V'(0) = a$ as the initial condition for V', where a is just a guess. Next, integrate the ODE until $x = 1$. If $V'(1) = 0$, then we guessed right. Usually, $V'(1)$ will not satisfy this desired condition at $x = 1$; it will either undershoot or overshoot. The idea is to find a value of a so that $V'(1) > 0$ and a different value so that $V'(1) < 0$ and then choose a between these two values, continuing to iterate until convergence is obtained.

Another common way to solve BVPs is to divide the interval [here $(0, 1)$] into m nodes and assign a value of V to each of these nodes. This will result in a division of the BVP into a large set of algebraic equations. These are solved using an equation solver. AUTO (see the software appendix) uses this method to solve BVPs. The advantage that this method has over shooting is that it is not sensitive to the stability of the equation's solution.

Noisy ODEs

A common model for noisy neurons adds normally distributed noise to the right-hand sides of an equation, resulting in a stochastic differential equation (SDE):

$$du = G(u, t)dt + dW(t). \tag{1.49}$$

Here dW is the derivative of Brownian motion, also called a Wiener process. We won't formally say anything more about these fascinating models as they are well beyond the scope of this chapter. Rather, we show how to simulate these equations. Suppose that $dW(t)$ has a zero mean and a variance σ^2. Then the Euler method for simulating this equation has the form:

$$u_{n+1} = u_n + hG(u_n, t_n) + \sigma\sqrt{h}N(1), \tag{1.50}$$

where $N(1)$ is a normally distributed random number with unit variance and zero mean. The unusual part of this model is that the square root of the step size is used to multiply the noise. The effective variance is then $\sigma^2 h$, which scales linearly with the step size.

Another source of noise is the random flipping of channels. Consider the simple process of a channel flipping between two states:

$$C \to O$$

$$O \to C$$

where O is open and C is closed. The rates are, respectively, α and β. If the channel is in the closed state, then the probability that it opens in time h is determined by a Poisson process and is approximately αh if h is small. We must choose h small enough so that during the same time the chance that it flipped back is essentially zero. Thus the numerical scheme for this type of noise is to pick a random number between 0 and 1. If this number is less than αh, then flip the channel to the open state; otherwise keep it closed. If there are many channels, then this procedure must be applied to each of them independently. This scheme only approximates the dynamics and gets more and more accurate as h gets smaller. We will call this the *Markov algorithm*.

An exact simulation can be obtained by using the *Gillespie algorithm* (see chapter 3, section 3). Let N be the number of closed channels and let M be the number of open channels. Let $r = \alpha N + \beta M$. r is the rate at which some event will happen; that is, one of the N closed channels opens or the M open channel closes. We generate a random time for the next event with this rate:

$$t_{\text{next}} = -\frac{1}{r} \ln R_1, \tag{1.51}$$

where R_1 is a random number between 0 and 1. This tells us when the next event will occur. Next we need to figure out which event. Let R_2 be a random number between 0 and r. If $R_2 < \alpha N$, then the event is a closed-to-open transition; otherwise it is an open-to-closed transition. In the former case, decrease N and increase M, while in the latter case do the opposite.

The Gillespie algorithm is exact, so that it is always better than the Markov algorithm. However, if there are only a few channels or the rates are very slow, then the time between events can be very large. If the stochastic system is coupled to some set of ordinary differential equations, then the solution to the ODE between events needs to be computed and using the step size between events can be very inaccurate. Thus, if the rates are very slow or there are very few channels, it is better to use the Markov algorithm, which has a fixed step size. For additional comparisons of algorithms in this context, see Mino et al. (2002).

PDEs

The easiest way to simulate a PDE as in figure 1.3 is to discretize it to a set of ODEs. Consider the passive cable (with time constant, τ and length constant, λ) of length l (see chapter 10, section 1):

$$\tau \frac{\partial V}{\partial t} = -V(x,t) + \lambda^2 \frac{\partial^2 V}{\partial x^2}$$

$$V(0,t) = V_a \tag{1.52}$$

$$\frac{\partial V(l,t)}{\partial x} = 0.$$

We consider the potential at $N + 1$ discrete location spaces an amount $k = l/N$ apart, and define $V_j(t) = V(j\Delta x, t)$ where $j = 0, \dots, N$. We approximate the spatial derivative by the difference:

$$\frac{\partial^2 V}{\partial x^2} \approx \frac{V_{j+1} - 2V_j + V_{j-1}}{\Delta x^2} \tag{1.53}$$

and obtain the set of ODEs:

$$\tau V_j' = -V_j + (\lambda^2/\Delta x^2)(V_{j+1} - 2V_j + V_{j-1}). \tag{1.54}$$

At $j = 0$ and $j = N$, we need to take the boundary conditions into account. At $j = 0$, we specified the voltage to be $V(0,t) = V_a$, so that the $j = 1$ equation is

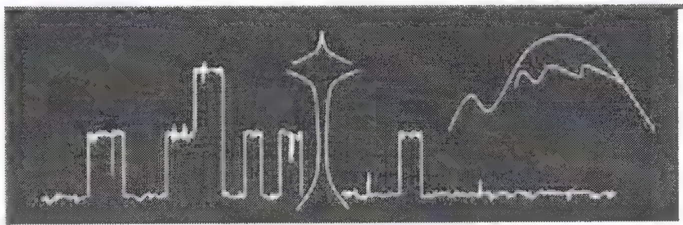
$$\tau V_1' = -V_1 + (\lambda^2/\Delta x^2)(V_2 - 2V_1 + V_a). \tag{1.55}$$

At $x = l$, corresponding to $j = N$, we have the zero-flux boundary condition, which is equivalent to imposing even symmetry about $x = l$. Even symmetry means $V_{N+1} = V_{N-1}$ and then substituting into the differential equation at $j = N$ we get:

$$\tau V_N' = -V_N + 2(\lambda^2/\Delta x^2)(V_{N-1} - V_N). \tag{1.56}$$

This use of even symmetry represents the zero-flux boundary condition more accurately than the one-sided differences for the boundary conditions that we used in the simulation for figure 1.4. Now, we solve this system using our favorite ODE solver. We remark that if N is large, then the step size Δx is small and the parameter $\lambda^2/\Delta x^2$ can be very large. Thus, stiff ODE solvers are often used for PDEs that arise from the cable equation.

Acknowledgments
Bard Ermentrout is supported in part the National Institute of Mental Health and the National Science Foundation. John Rinzel is supported in part by National Institute on Deafness and Other Communication Disorders/National Institutes of Health, grant DC008543-01.



ROLE OF THE CORTICAL NEURON: INTEGRATOR OR COINCIDENCE DETECTOR?

MOSHE ABELES

Institute of Physiology, University of Lausanne, Lausanne, Switzerland

ABSTRACT. A model was constructed of the relations between the incoming excitatory postsynaptic potentials and inhibitory postsynaptic potentials to the intracellular membrane potential fluctuations and to the firing rate of a single neuron. From this model, the strength of synapses was assessed in two ways: 1) the ability of several synchronous presynaptic spikes to initiate a postsynaptic spike, termed the synchronous gain of the synapse; and 2) the ability of several asynchronous presynaptic spikes to add a spike to the output spike train, called the asynchronous gain. It was found that for the conditions prevailing in the brain's cortex, the synchronous gain is almost always higher. It appears that the cortical neurons act as coincidence detectors, and that the appropriate code for the higher cortical functions is coincidence and not firing rate.

Isr J Med Sci 18: 83–92, 1982

The biophysics of neural coding

Adrienne Fairhall
Department of Physiology and Biophysics
University of Washington, Seattle

From single neuron to adaptive feature detector

What can a neuron compute?

Integrators or differentiators?
How do we quantify computation?
What kinds of behavior are possible from single neurons?

How is the computation established by the ion channel makeup?

Neurons are diverse.
What is the relationship between the dynamical ion channel description and a functional computational model?

To what extent do single neurons show adaptive behavior?

Fast adaptation to stimulus statistics may be a product of intrinsic neural nonlinearities
What a neuron represents depends both on its composition and on the statistics of the stimulus

From single neuron to adaptive feature detector

Experimental/empirical approach

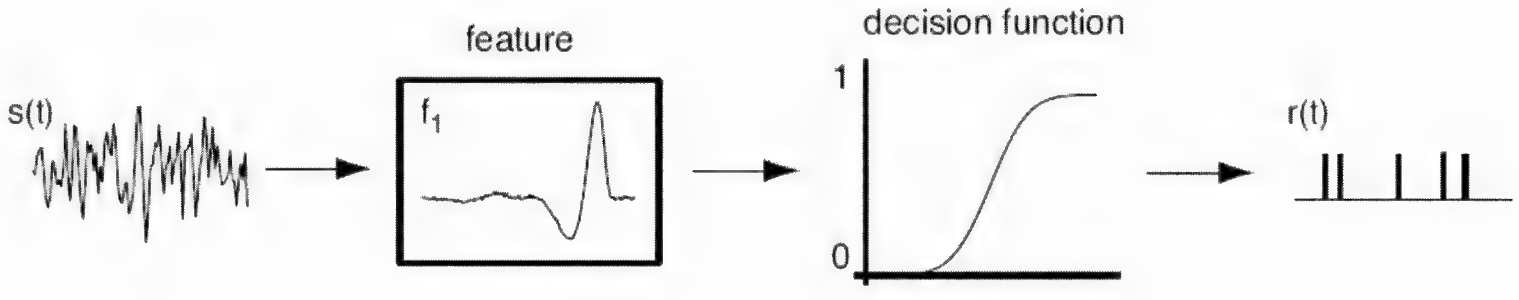
- White noise methods
- Applied to model neurons
- Applied to auditory neurons

Analytical approach

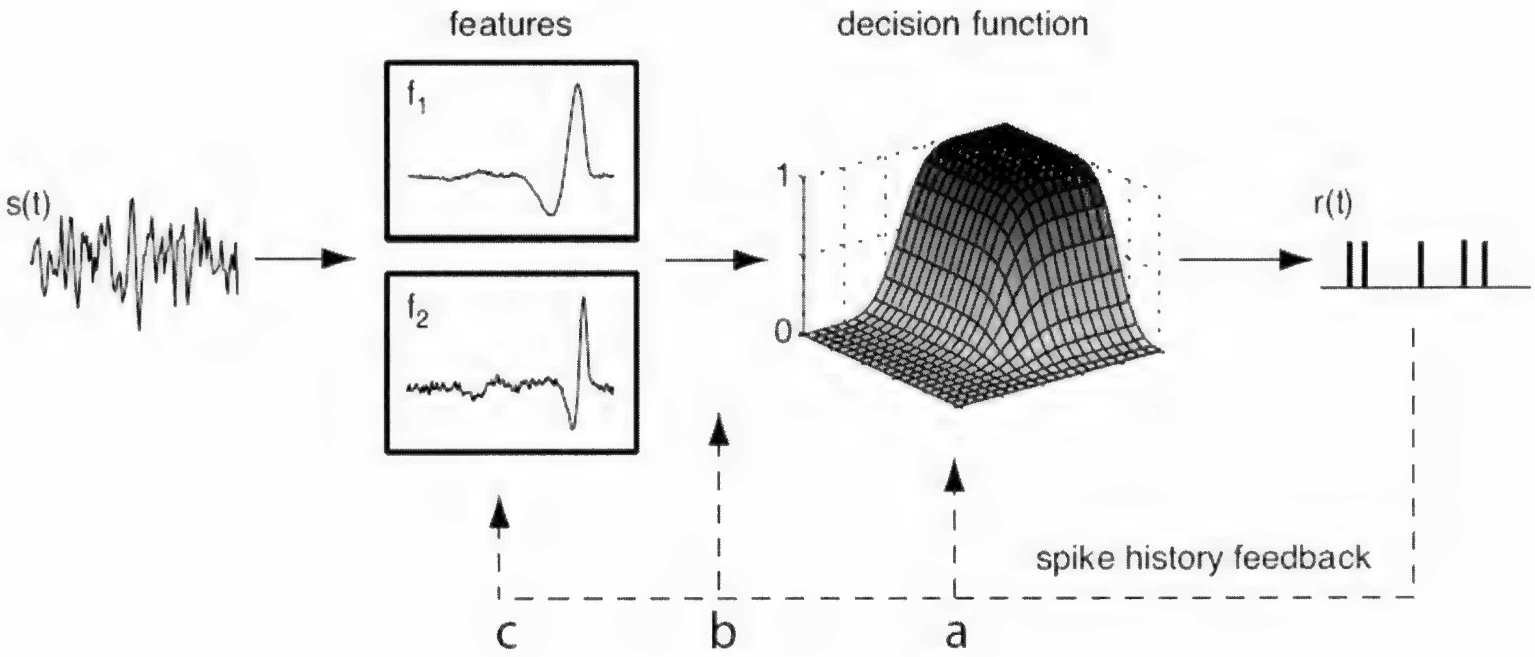
solving for equivalent linear/nonlinear models
starting from the dynamical system

Implications for adaptation

Neuronal representation of information


s(t)
feature
f1
decision function
1
0
r(t)

Neuronal representation of information


features
decision function
s(t)
f1
f2
1
0
r(t)
spike history feedback
c
b
a

$$C\frac{dV}{dt} = I(t) - \bar{g}_{\mathrm{K}} n^4 (V - V_{\mathrm{K}}) - \bar{g}_{\mathrm{Na}} m^3 h (V - V_{\mathrm{Na}}) - \bar{g}_l (V - V_l)$$

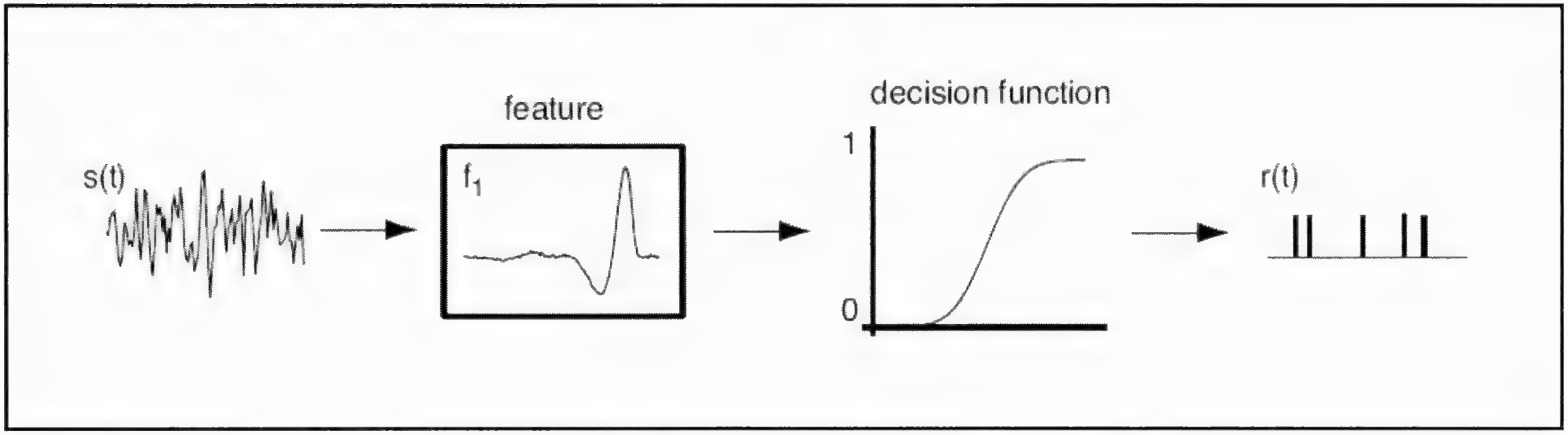
$$\begin{aligned} dn/dt &= \alpha_n(V)(1-n) - \beta_n(V)n \\ dm/dt &= \alpha_m(V)(1-m) - \beta_m(V)m \\ dh/dt &= \alpha_h(V)(1-h) - \beta_h(V)h \end{aligned}$$

$$\begin{aligned} \alpha_n(V) &= \frac{10 - V}{100\,(\exp((10-V)/10) - 1)} \\ \alpha_m(V) &= \frac{25 - V}{10\,(\exp((25-V)/10) - 1)} \\ \alpha_h(V) &= 0.07\exp(-V/20) \end{aligned}$$

$$C = 1\,\mu\mathrm{F/cm}^2$$

$$\begin{aligned} \bar{g}_{\mathrm{K}} &= 36\,\mathrm{mS/cm}^2 & V_{\mathrm{K}} &= -12\,\mathrm{mV} \\ \bar{g}_{\mathrm{Na}} &= 120\,\mathrm{mS/cm}^2 & V_{\mathrm{Na}} &= +115\,\mathrm{mV} \\ \bar{g}_l &= 0.3\,\mathrm{mS/cm}^2 & V_l &= +10.613\,\mathrm{mV} \end{aligned}$$

$$\begin{aligned} \beta_n(V) &= 0.125\exp(-V/80) \\ \beta_m(V) &= 4\exp(-V/18) \\ \beta_h(V) &= \frac{1}{\exp((30-V)/10) + 1} \end{aligned}$$



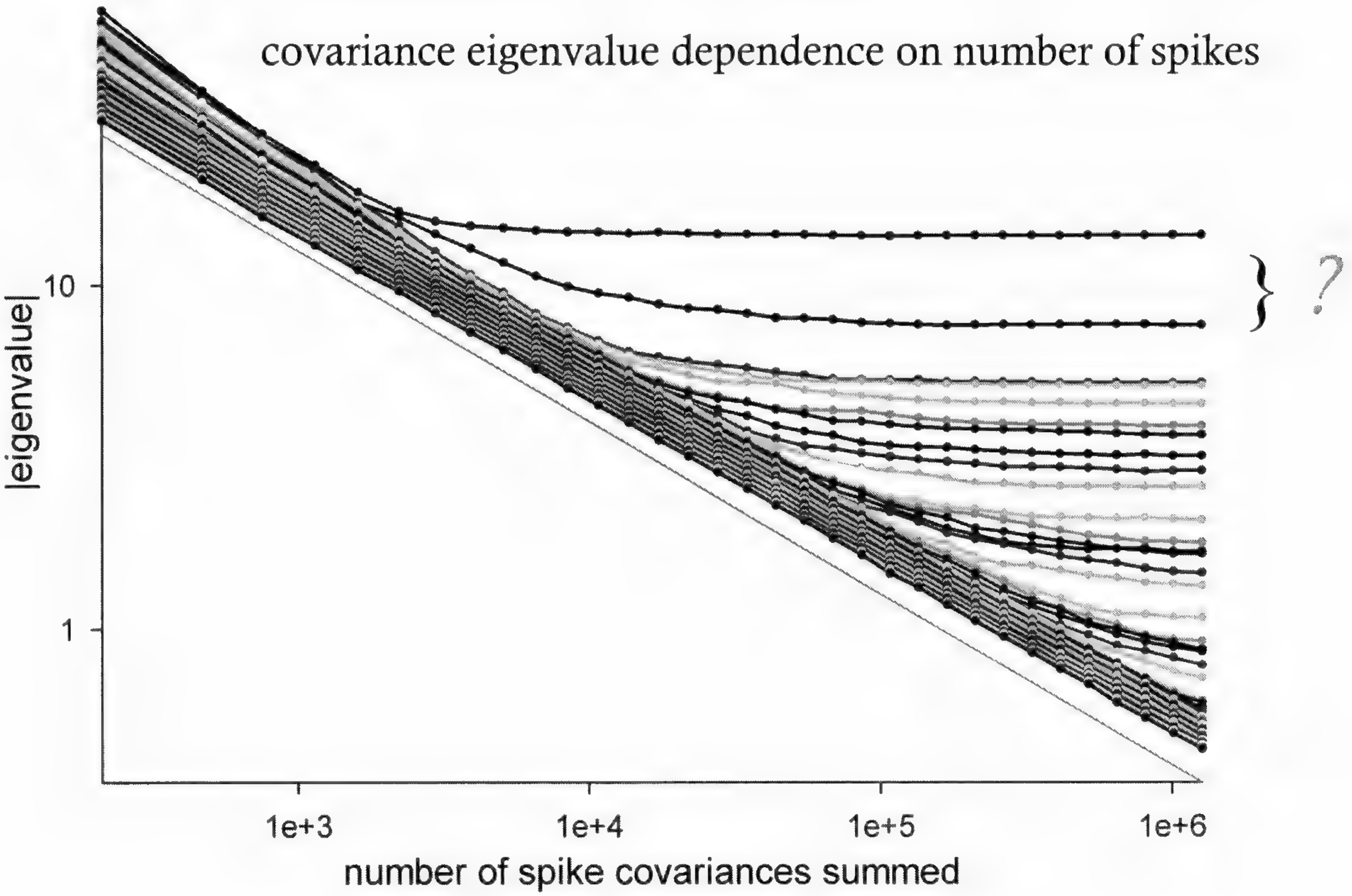
Dynamics to coding

Pro and con: method is “black box” approach for determining a general model of the system.

Can we break open the box for the case of single neurons?

Can this give us some insight into mechanisms of gain control?

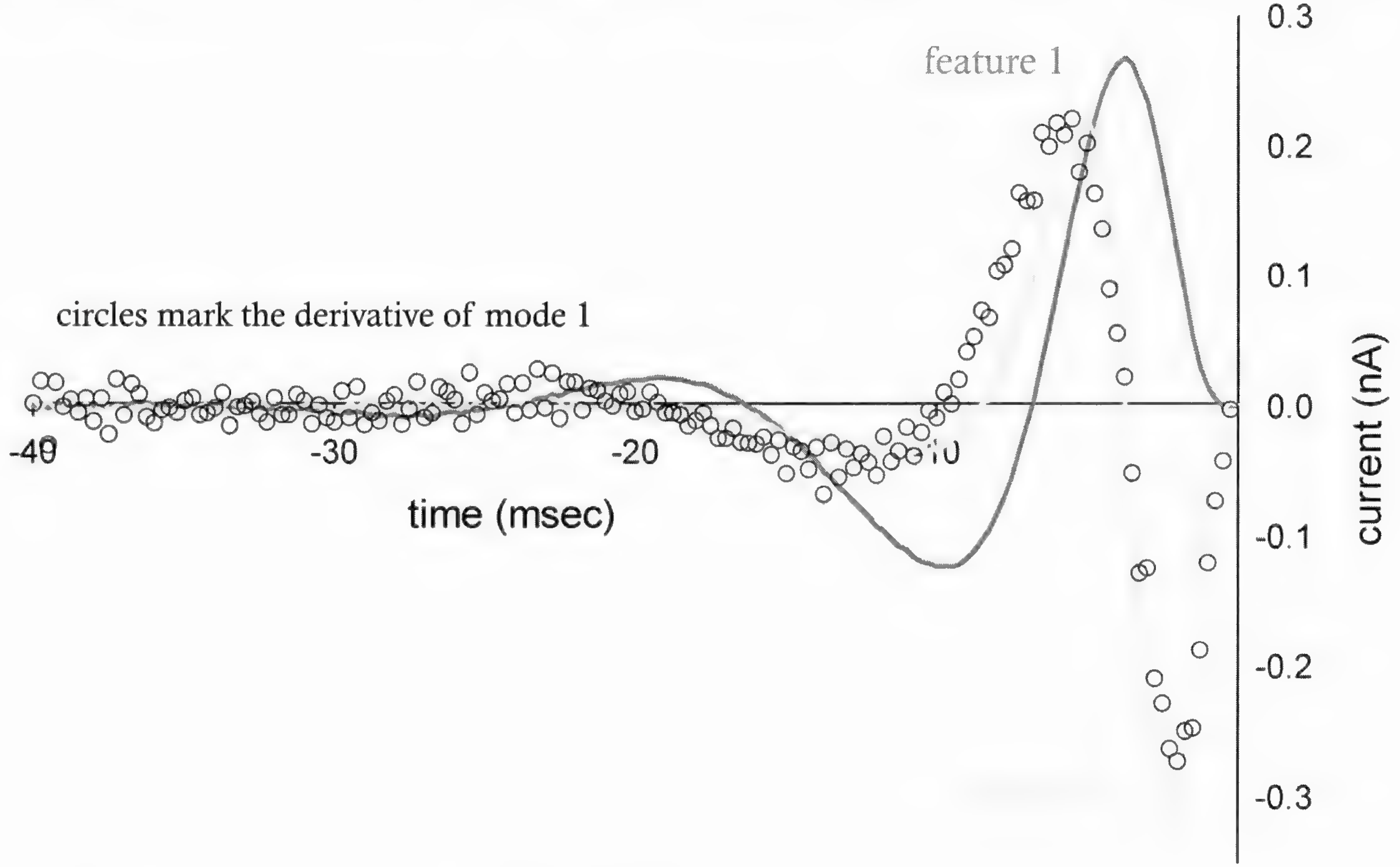
What kind of neuron is the Hodgkin-Huxley model?



covariance eigenvalue dependence on number of spikes

Blaise Aguera y Arcas and Bill Bialek

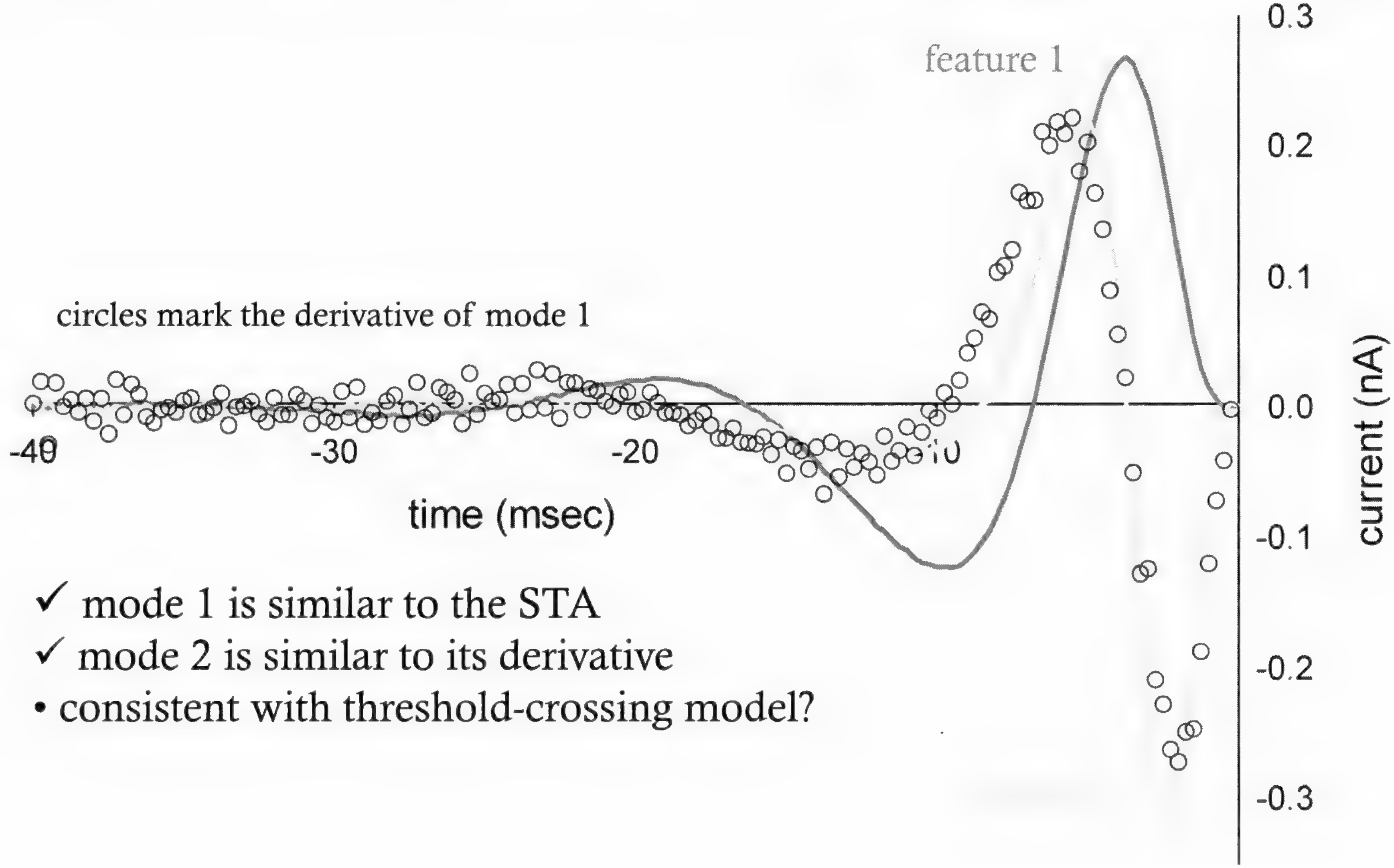
Most significant features for the Hodgkin-Huxley model neuron



circles mark the derivative of mode 1

Blaise Aguera y Arcas and Bill Bialek

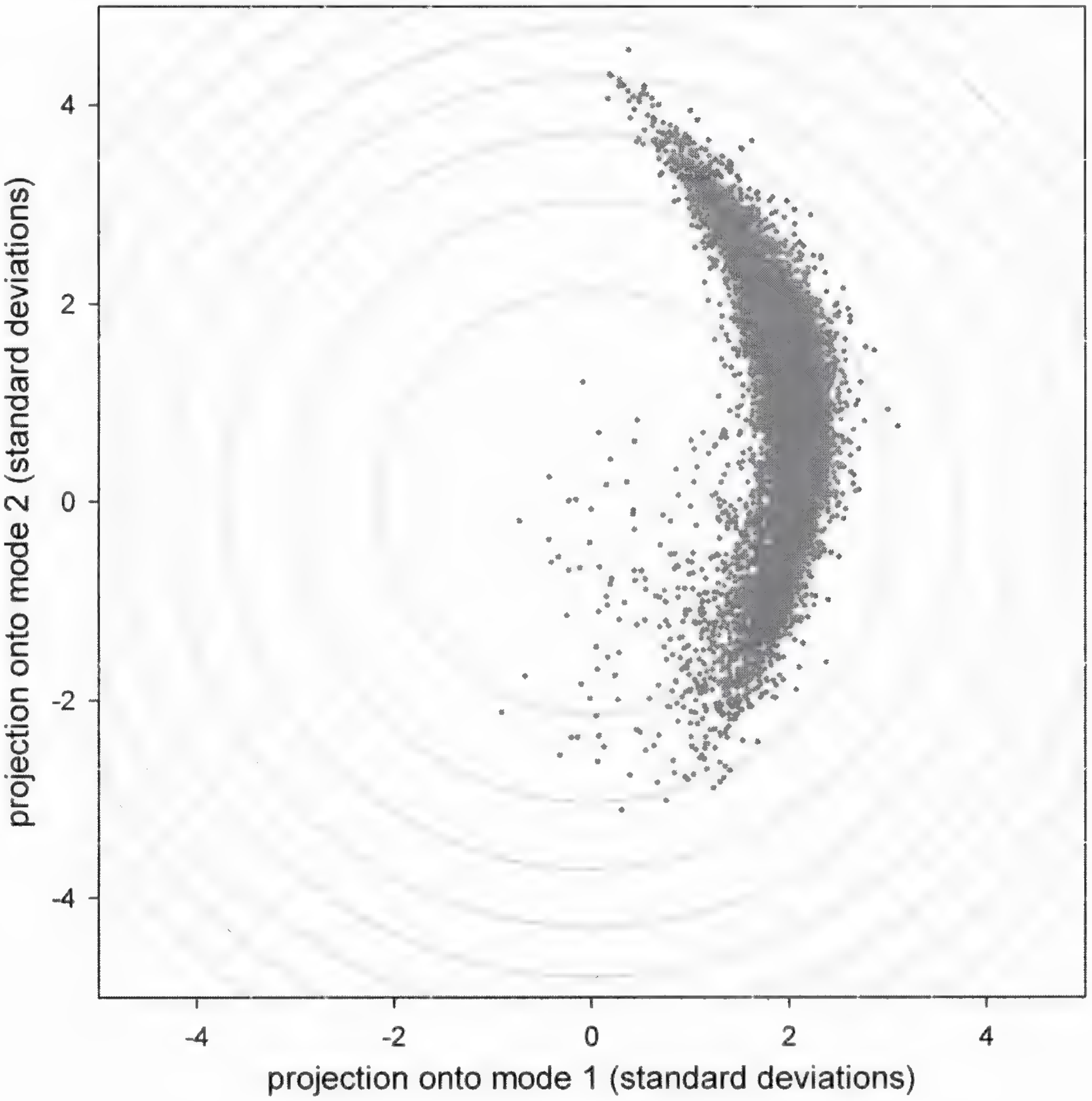
Most significant features for the Hodgkin-Huxley model neuron



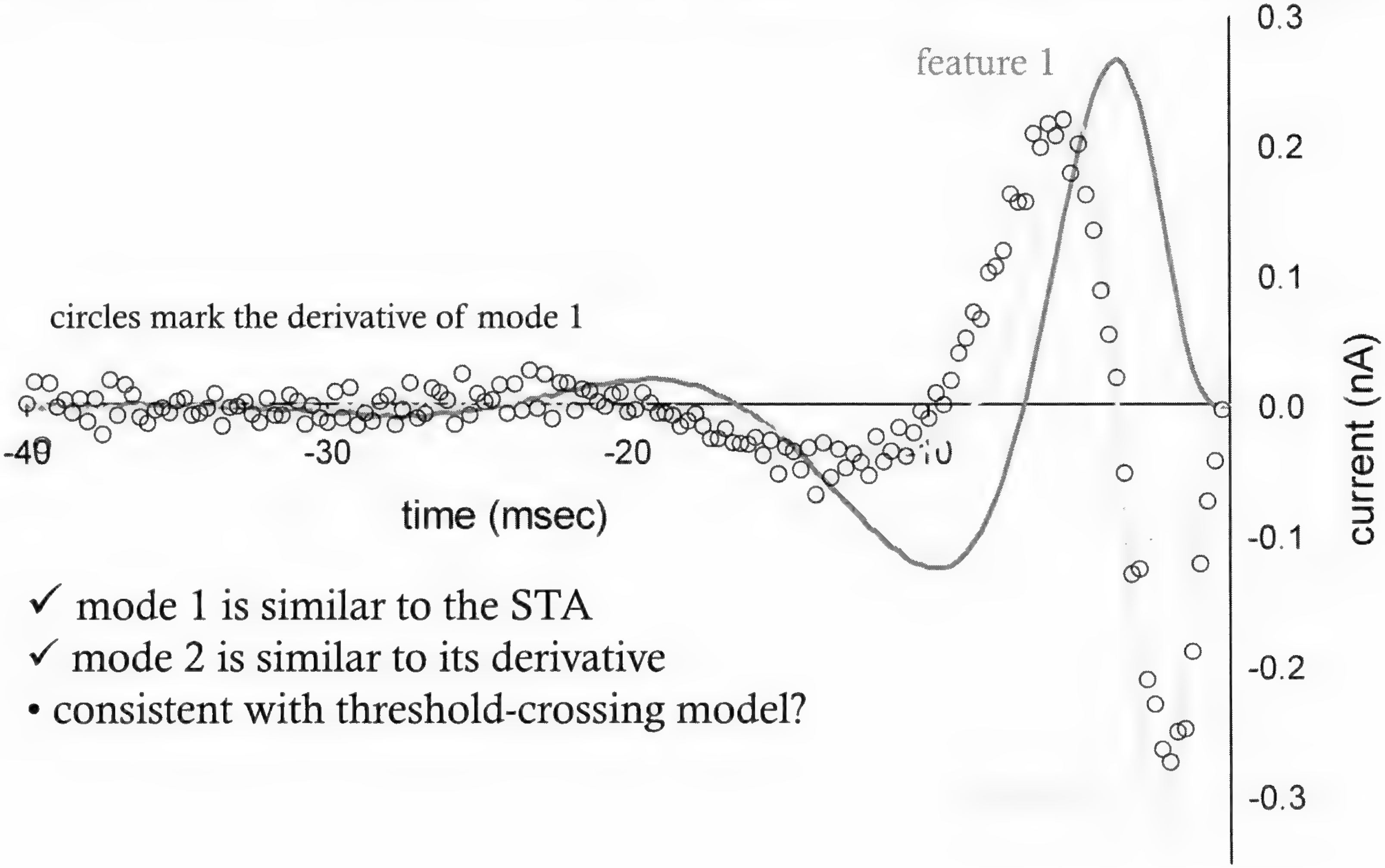
circles mark the derivative of mode 1

- ✓ mode 1 is similar to the STA
- ✓ mode 2 is similar to its derivative
- consistent with threshold-crossing model?

Spike-triggering stimuli in the space of the first two features



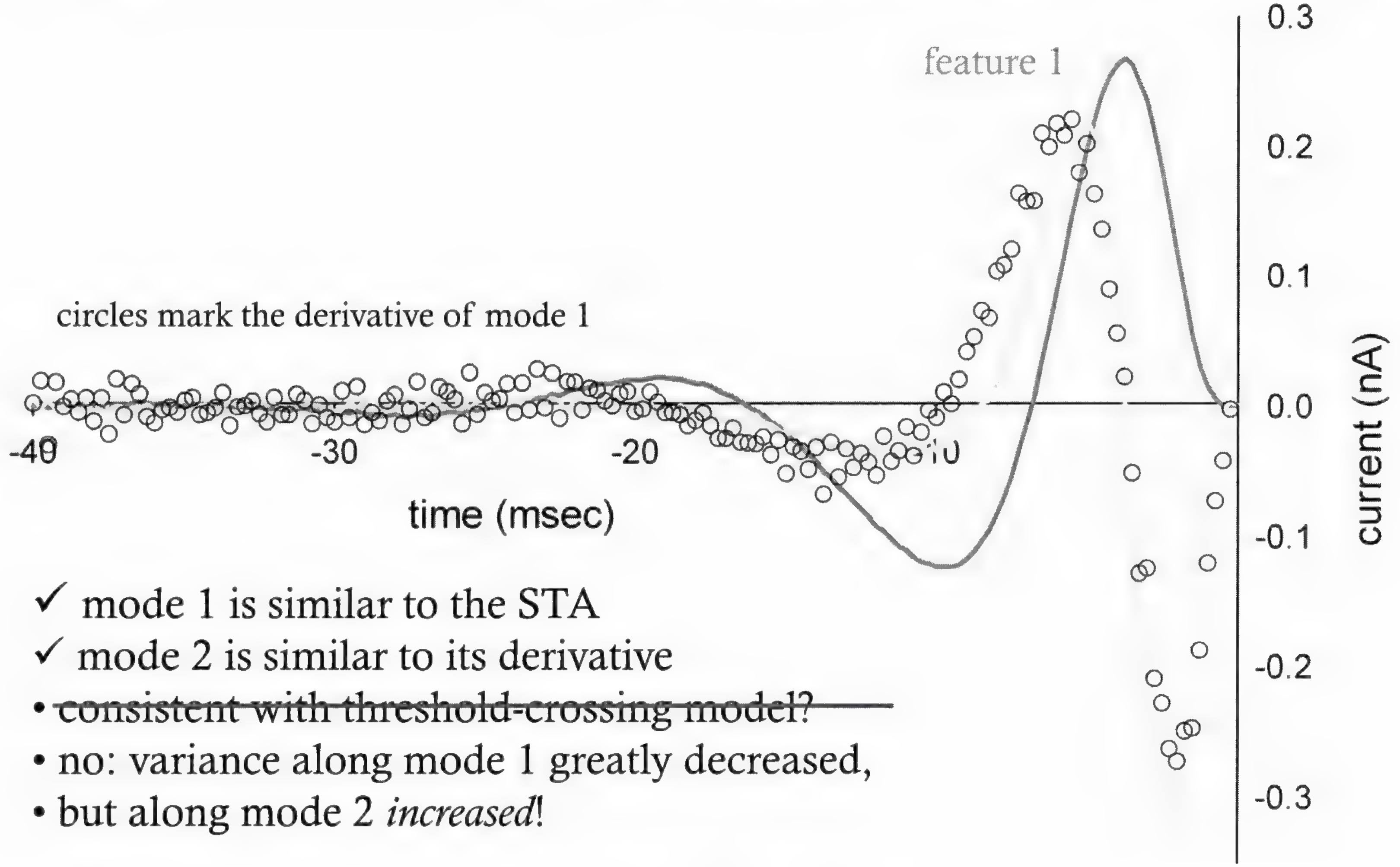
Most significant features for the Hodgkin-Huxley model neuron



circles mark the derivative of mode 1

- ✓ mode 1 is similar to the STA
- ✓ mode 2 is similar to its derivative
- consistent with threshold-crossing model?

Most significant features for the Hodgkin-Huxley model neuron

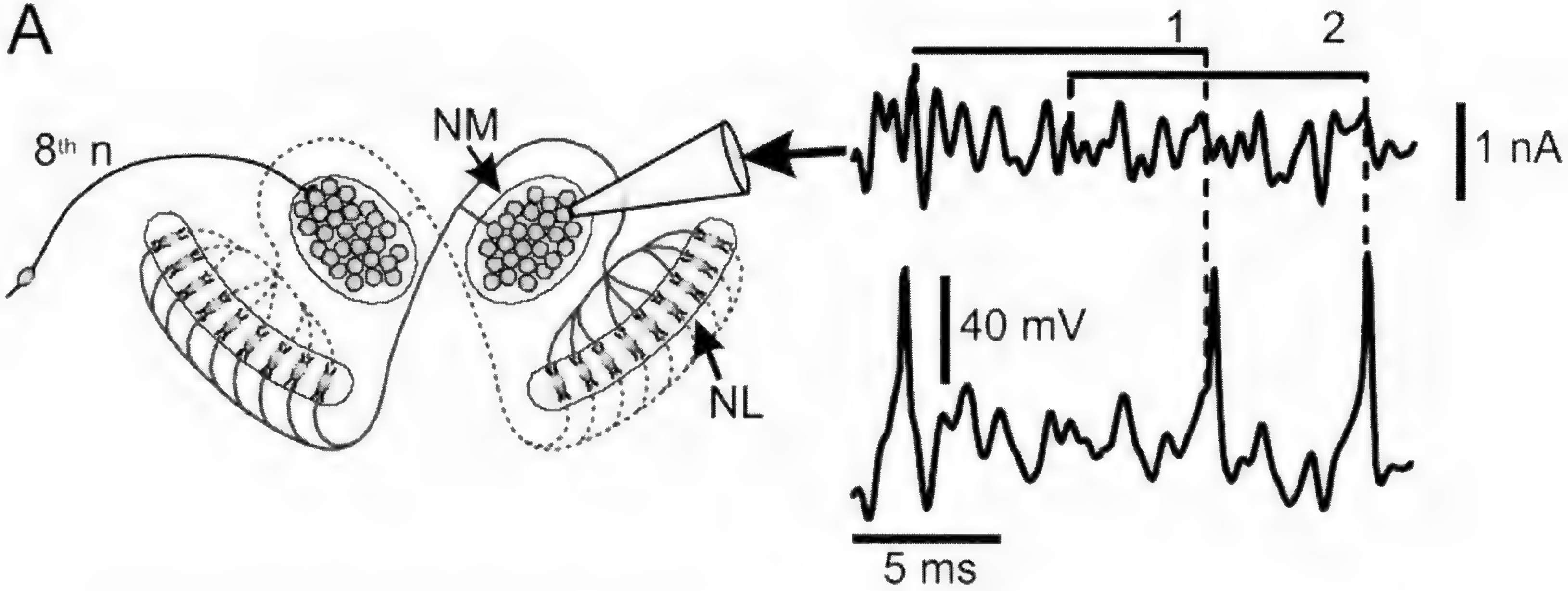


circles mark the derivative of mode 1

- ✓ mode 1 is similar to the STA
- ✓ mode 2 is similar to its derivative
- ~~consistent with threshold-crossing model?~~
- no: variance along mode 1 greatly decreased,
- but along mode 2 *increased*!

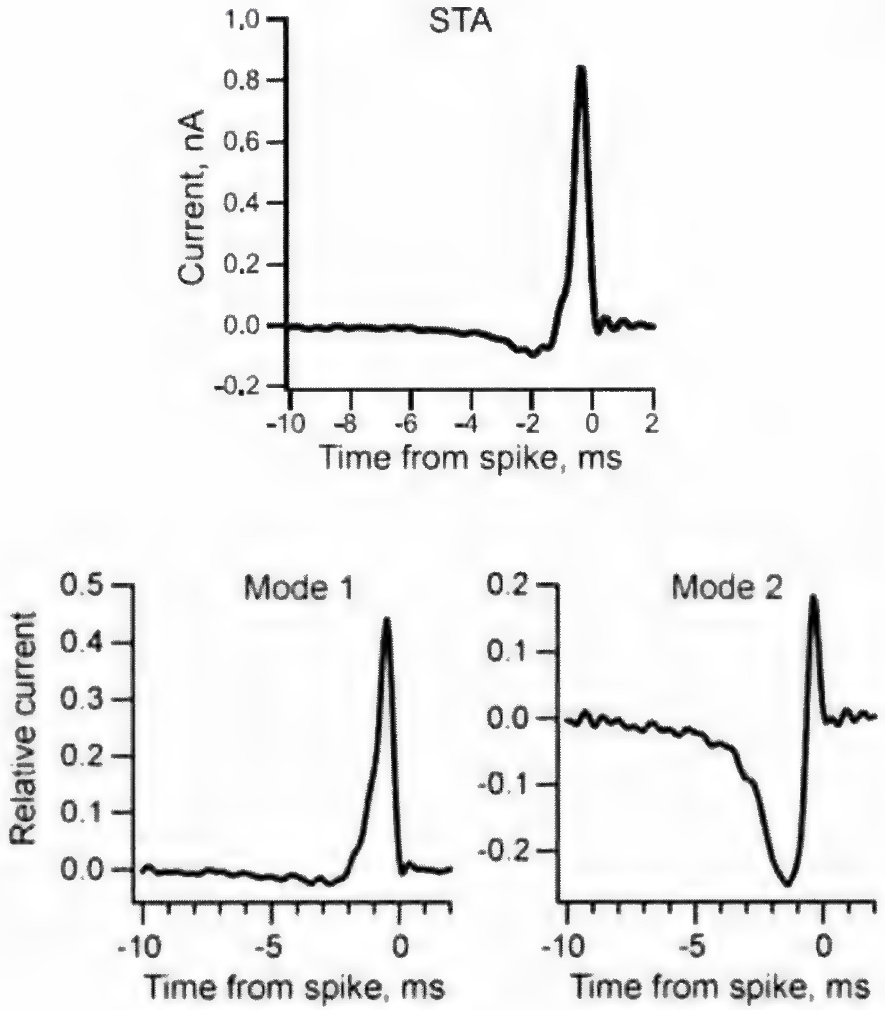
Differentiator-like, but not simple 1D filter-and-fire

Analysis of neurons in avian nucleus magnocellularis



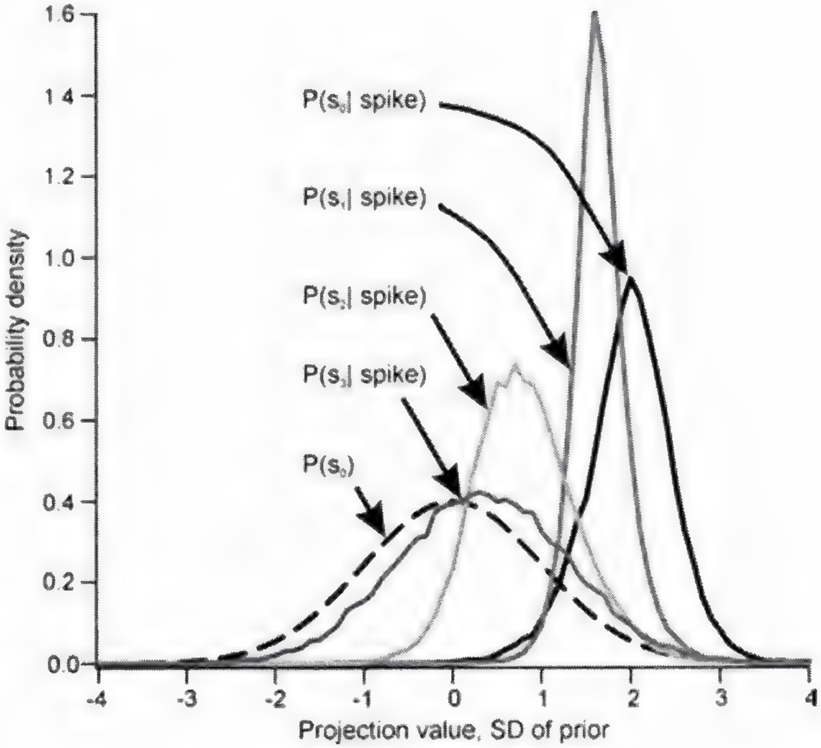
- Located in the avian brainstem
- Part of circuit for analysing temporal differences for sound localization
- inputs from auditory 8th nerve
- act as a DC filter to a current step
- given a time-varying input, what do they compute?

Sean Slee, Matt Higgs and Bill Spain Slee et al., J. Neurosci. (2005)



Determining the nonlinearity:

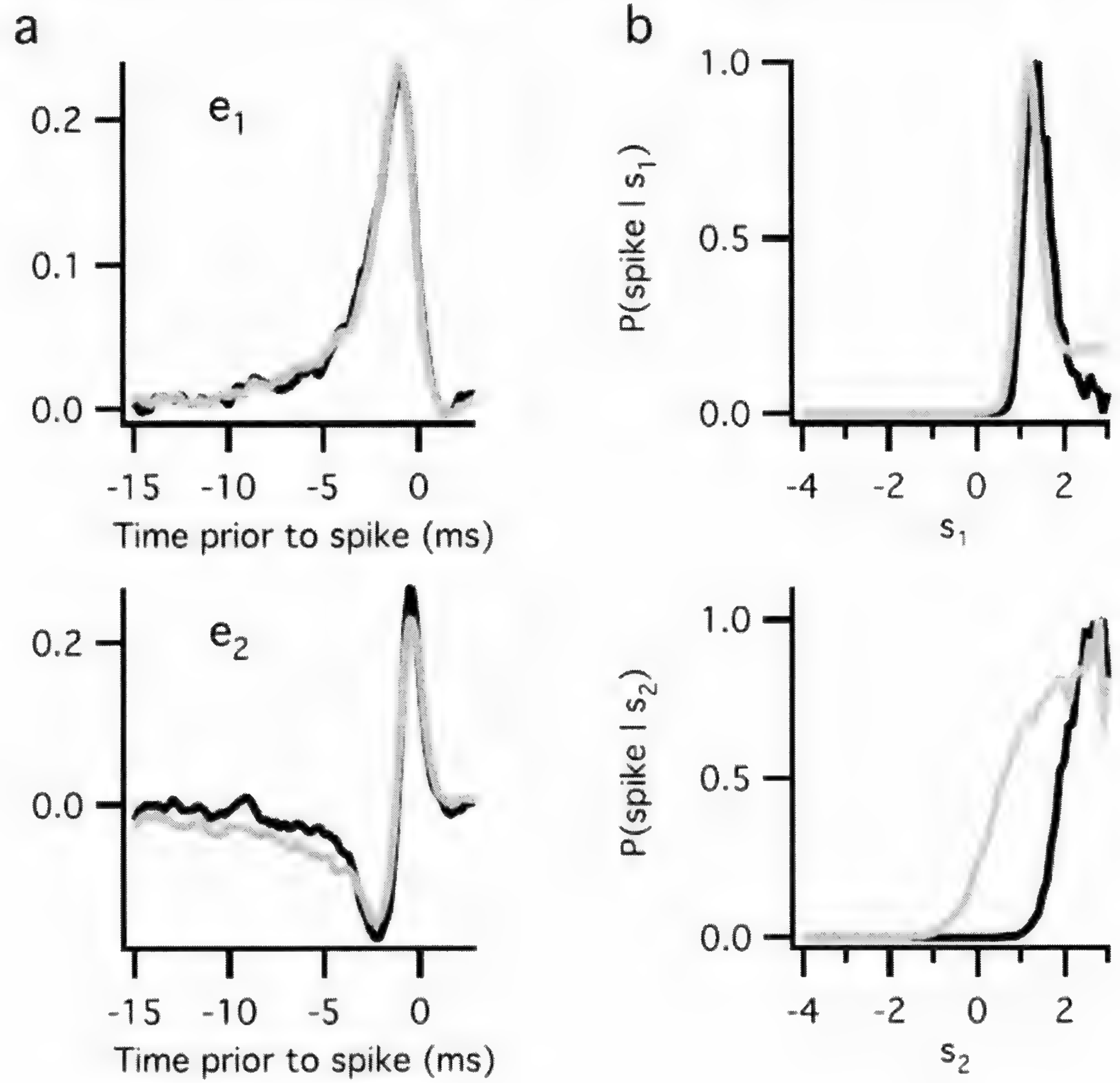
$$P(\text{spike}|s_1) = \frac{P(s_1|\text{spike})P(\text{spike})}{P(s_1)}$$



Dependence of the computation on channel kinetics

Treated the neurons with DTX, blocks Kv1.1, 1.2, 1.

Eliminates DC filtering



Examples of functional analysis of single neuron computation

The HH neuron is a differentiator to first order,
and has a threshold function in the space of mode 1 and mode 2
that is curved.

NM neurons have integrating and differentiating
components, and that their relative contributions were affected
by the presence of Kv channels

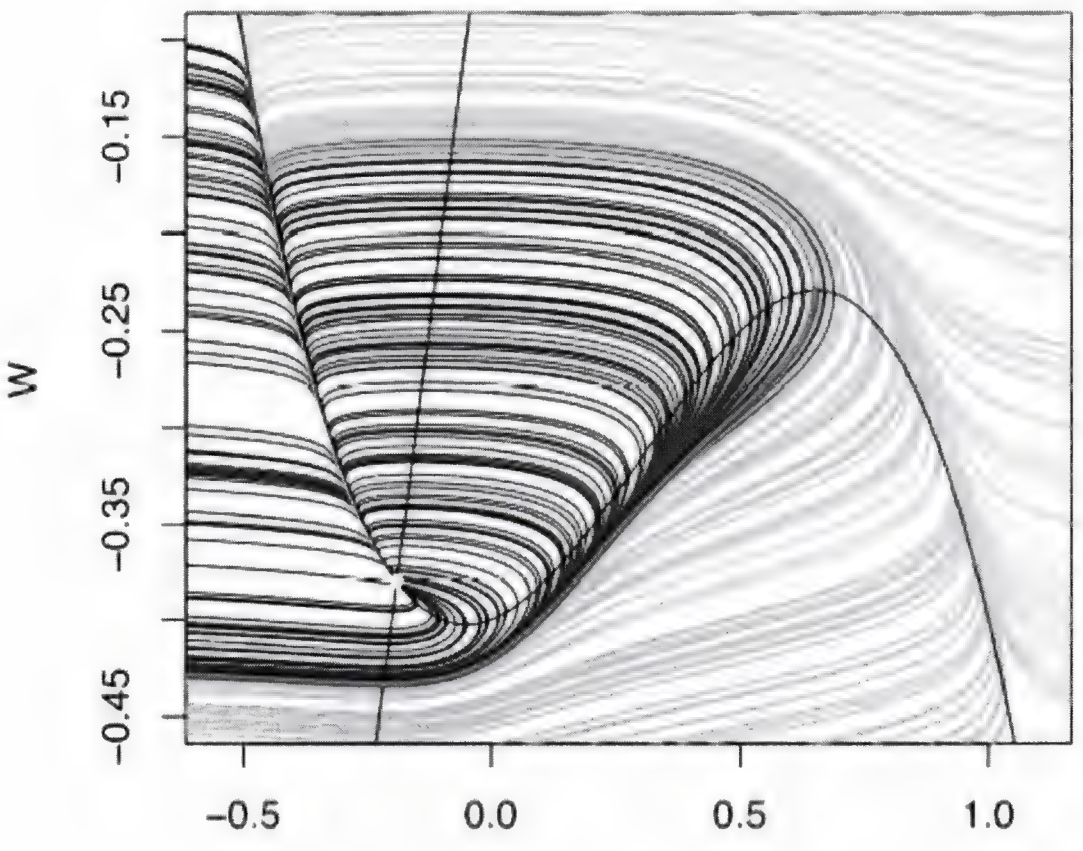
Can we get an analytic handle on the dependence of these features and the threshold function on channel properties?

Dynamics to coding

Deriving a linear/nonlinear model from the neuron's dynamics

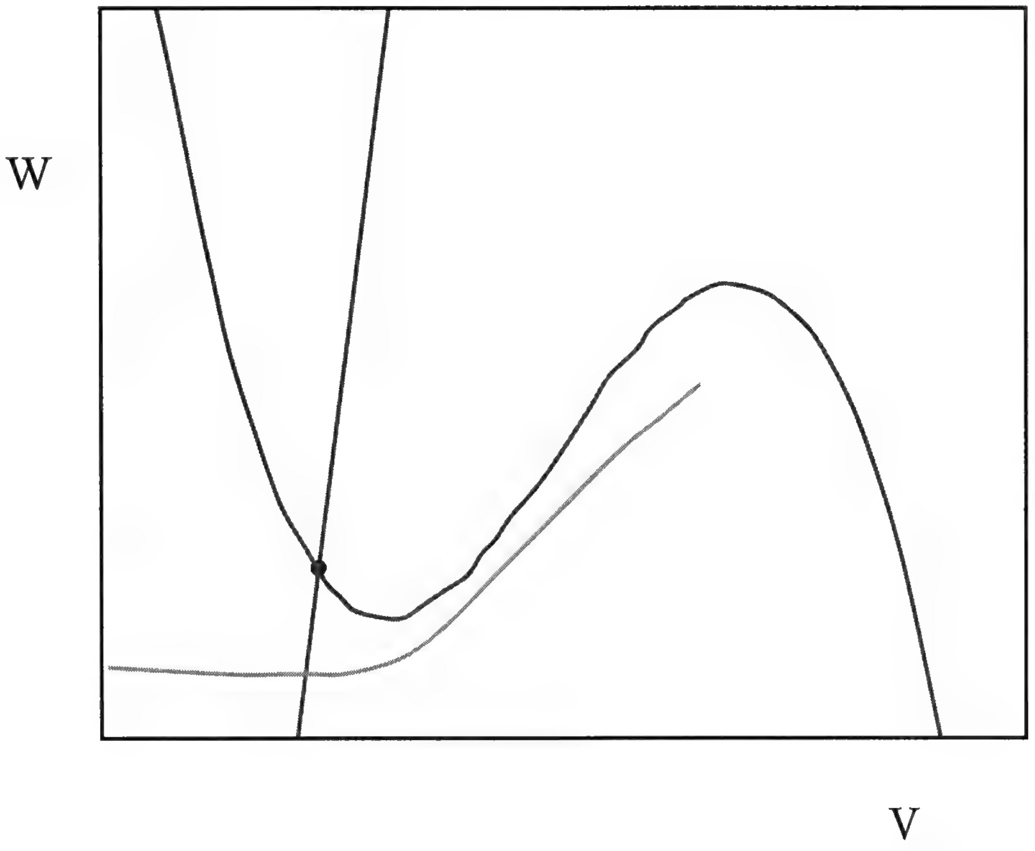
Sungho Hong
Blaise Aguera y Arcas
Mike Famulare

A simple model neuron: FitzHugh-Nagumo

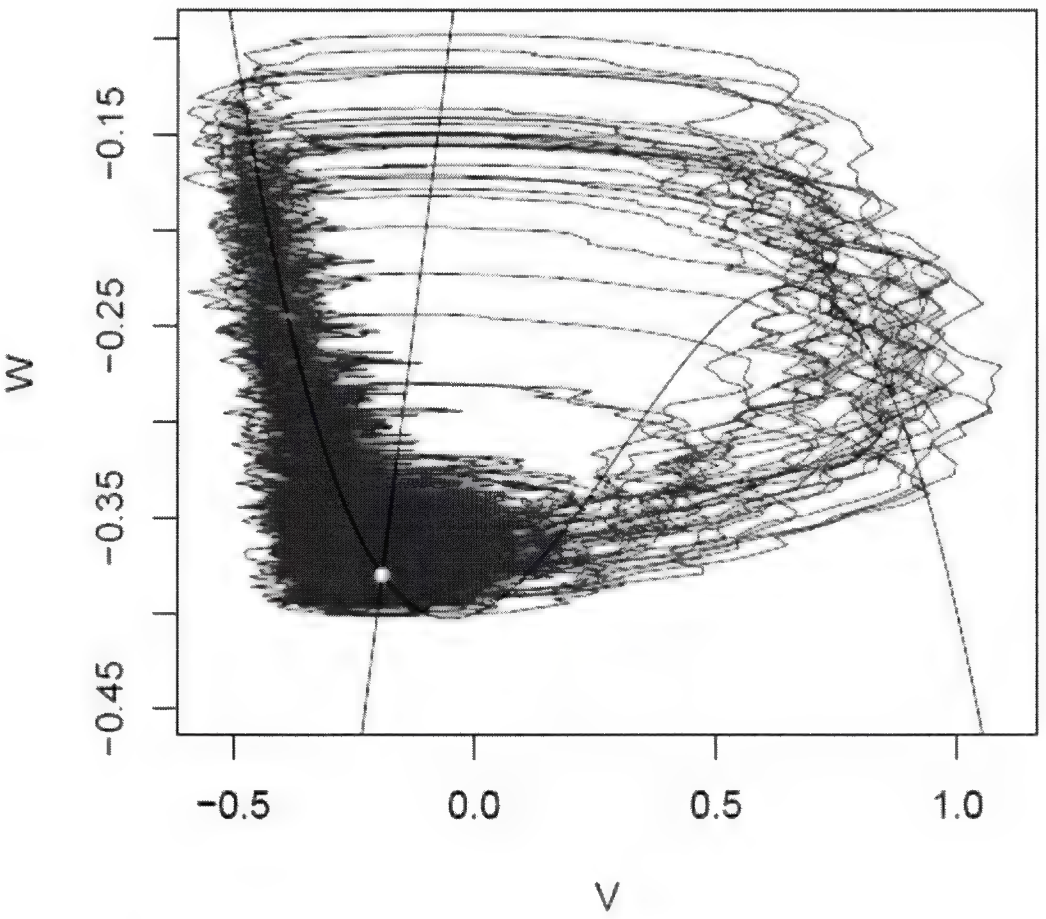


I = 0 phase portrait

Phase portrait of the FitzHugh-Nagumo neuron model



A noise-driven trajectory



Imagine I = 0 as determining the instantaneous flow if the current is suddenly turned off.

Linear analysis

General nonlinear first-order system:

$$\partial_t y_k = f_k(y_1, y_2, \ldots, y_N) + I(t)\delta_{k0}, \quad k = 0, 1, \cdots, n-1,$$

When the system has a stable fixed point: $y_k(t) = y_k^{(0)}$

we can linearize around the fixed point, and obtain a set of linear kernels.

→ These linear kernels have the form we are looking for: “features”.

$$\boxed{y_k^{(1)}(t) = \int dt' \, \epsilon_k^{(1)}(t-t')I(t').}$$

General forms of the features derive from the linearized system

Diagonalized solutions are of the form:

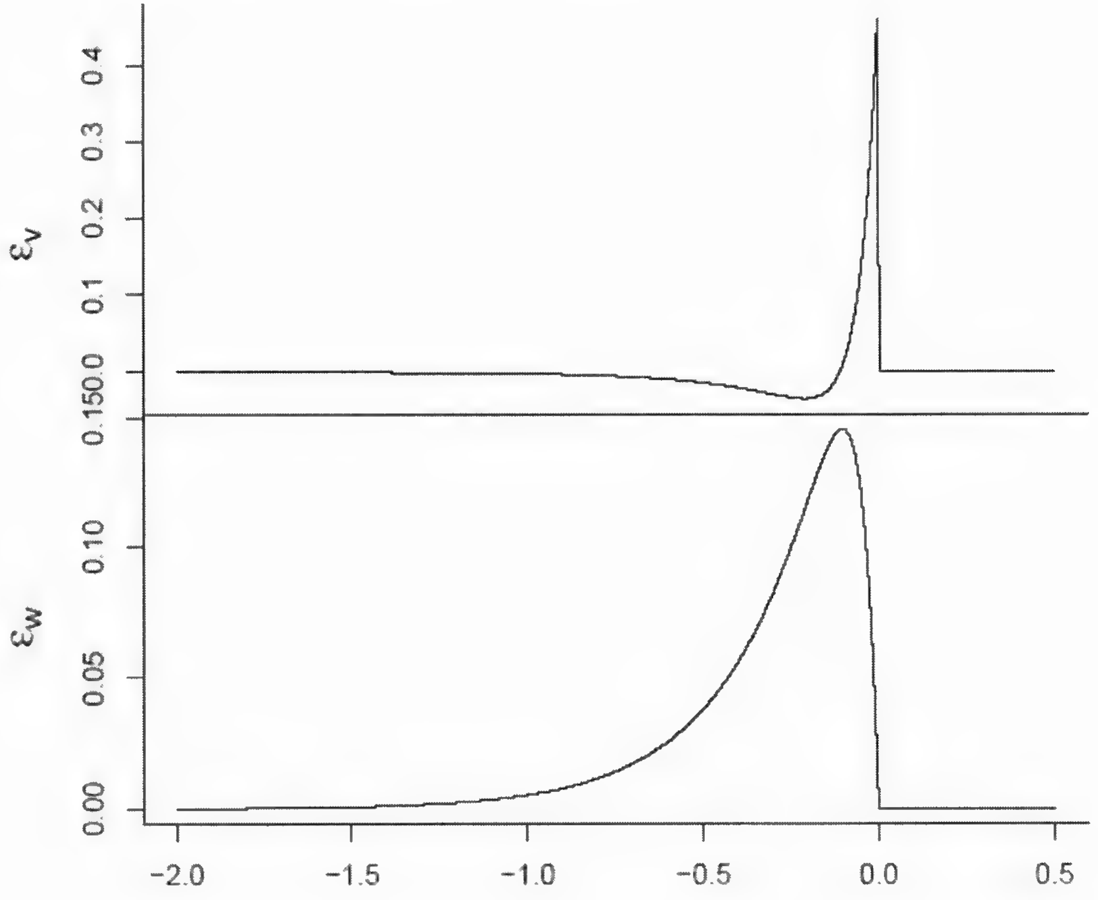
$$\tilde{\epsilon}_{\alpha}^{(1)}(t) = e^{\lambda_{\alpha} t}\theta(t)$$ where θ(t) is the Heaviside function

$$\boxed{\epsilon_i^{(1)}(t) = U_{i\alpha}^{-1}\tilde{\epsilon}_{\alpha}^{(1)}(t) = \sum^{\alpha} c_{\alpha} e^{\lambda_{\alpha} t}}$$

The kernels are simply sums of exponentials.

The rate constants result from derivatives of the original equations:
determined by the biophysics.

Linear filters for the FitzHugh-Nagumo system



General facts about the linear filters

When the linear approximation is reasonable,

→ features are identified as the linear kernels derived from the Jacobian of the system

→ there are therefore at most *n* features + *1* (for threshold crossing)

→ features all have a common form: sums of exponentials.

for purely real eigenvalues, purely decaying

leaky integrator-like

for eigenvalues with an imaginary part, oscillating

resonate-and-fire neurons

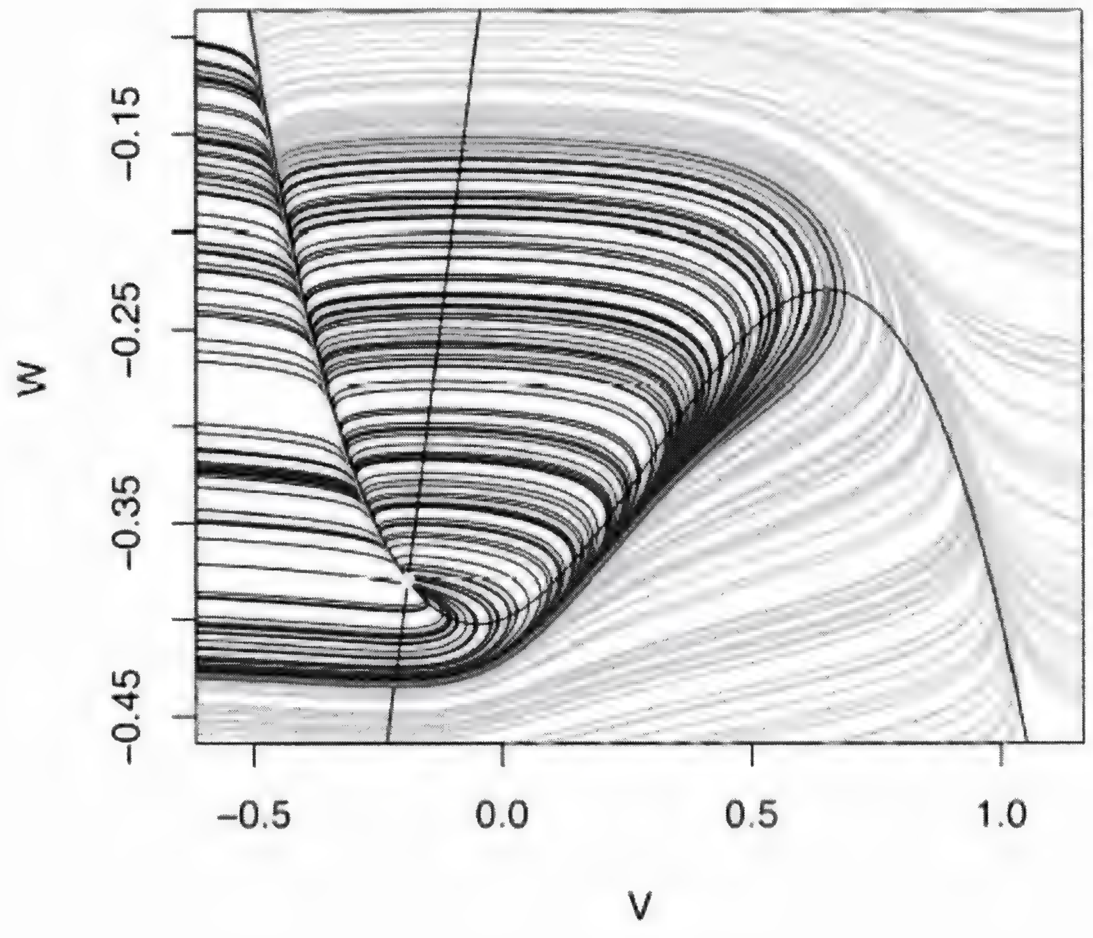
→ in principle, all features can be derived from the STA and its time derivatives

We aim to build a linear/nonlinear model.

This approximates the linear “front end” : how about the nonlinearity?

The threshold function

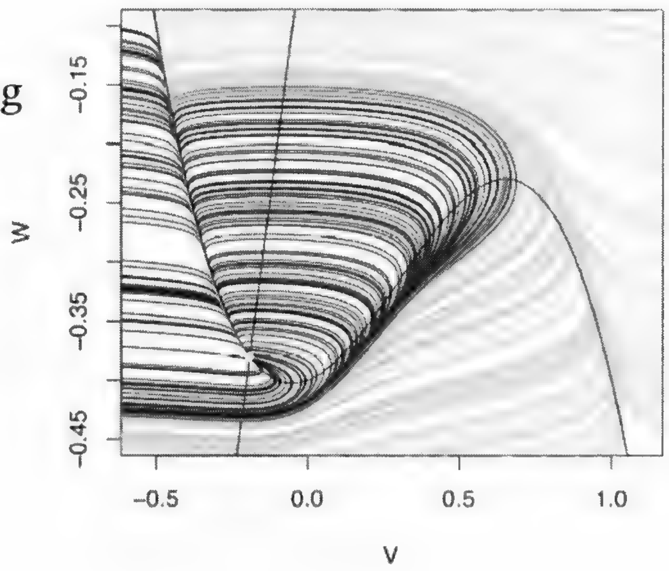
Typically, we define a spike as the crossing of a threshold in voltage



The threshold function

The phase plane shows that there is a boundary separating spiking from nonspiking trajectories.

This is the *threshold function:*
determine mathematically.



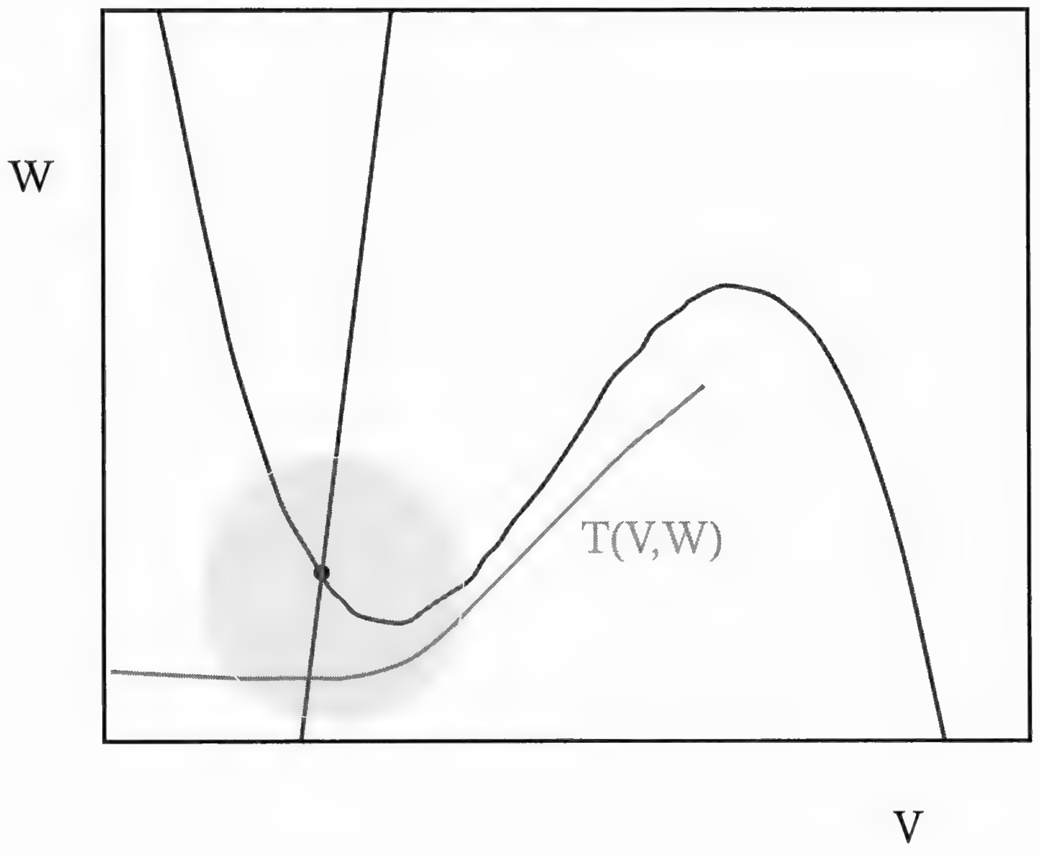
Define *relevant features* as the filters corresponding to dynamical variables that intersect with the threshold.

If the system is approximately linear up to threshold, there are at most n relevant features, where n is the number of dynamical variables. There can be fewer, $m < n$.

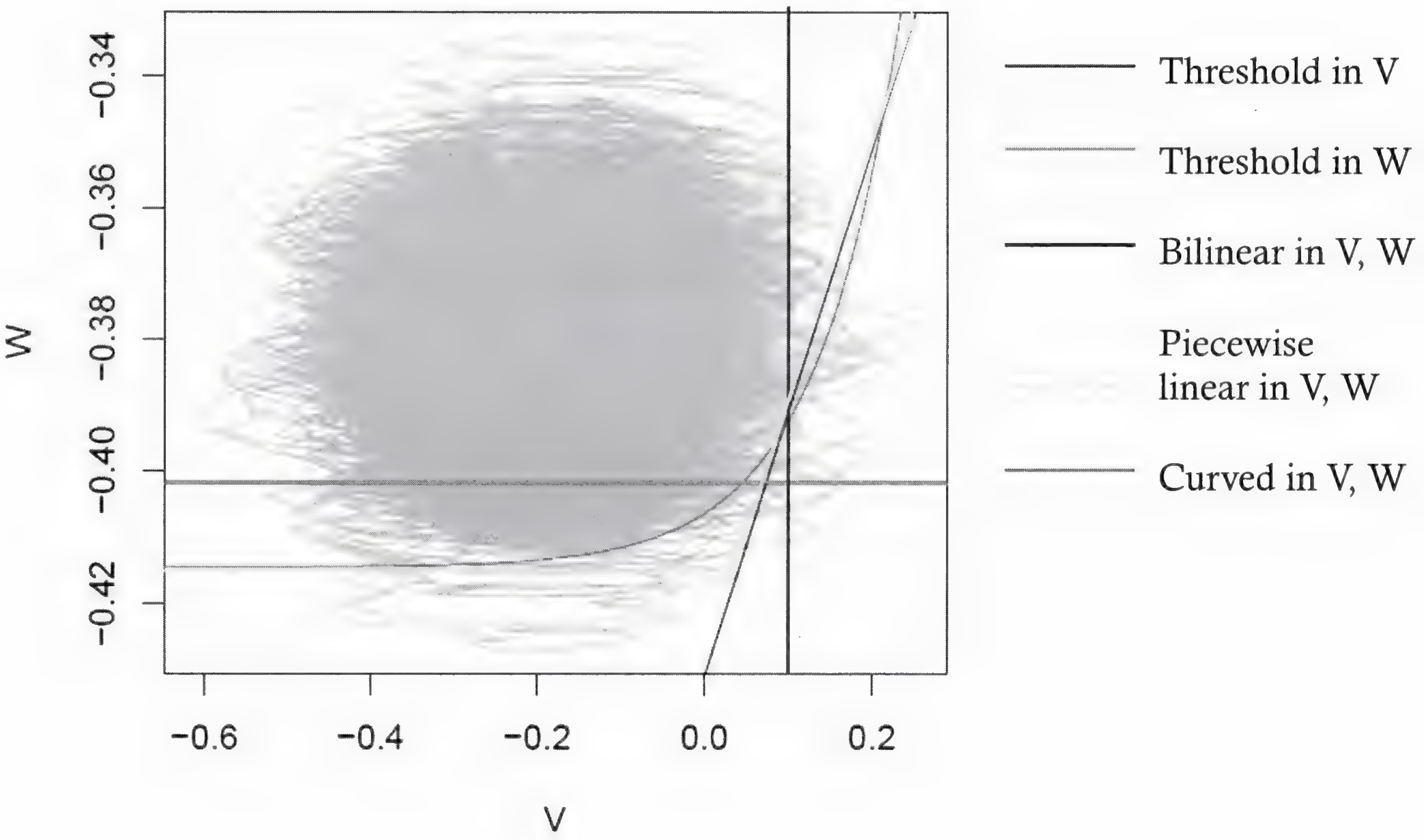
Generalization of the linear-to-threshold model:
m linear filters or m dimensions
the threshold is an m-dimensional curve

A multi-dimensional filter-and-fire model

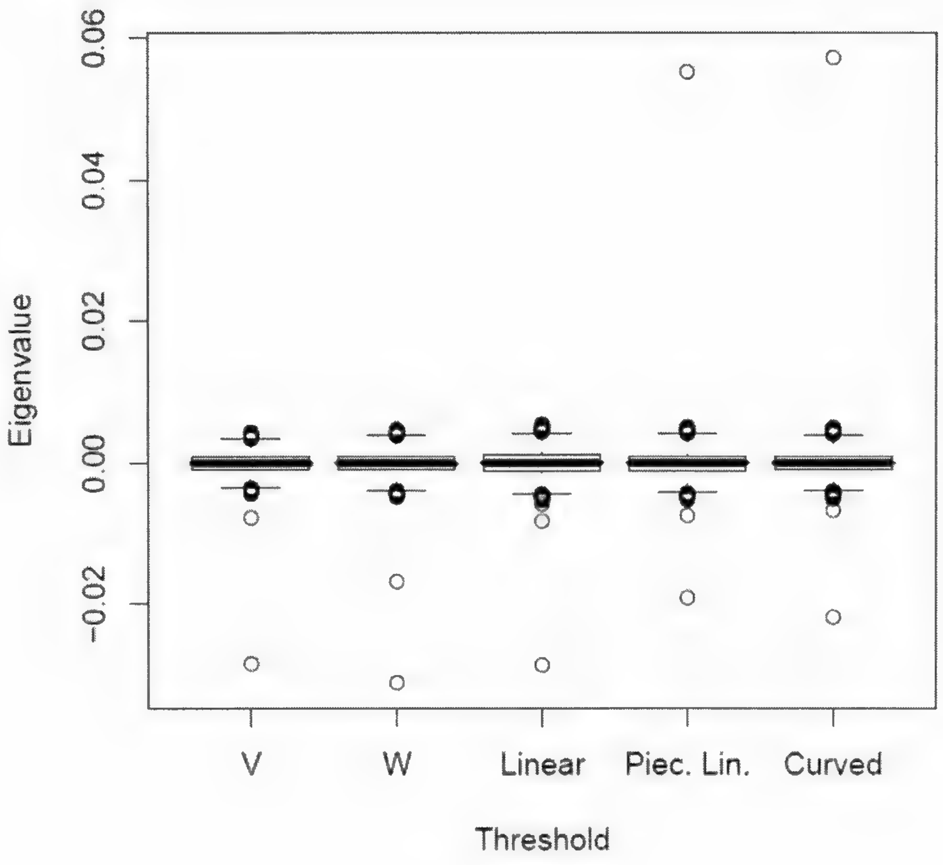
Driving with current I(t) moves you around in V according to $e_V * I$,
in W according to $e_W * I$,
until the state reaches the threshold, T(V,W)



A new model: filter-and-fire with a multidimensional threshold



Covariance analysis on the model

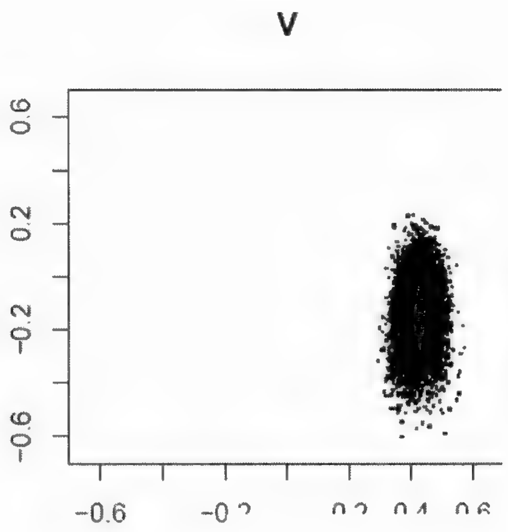


In each case, pick out
two relevant modes.

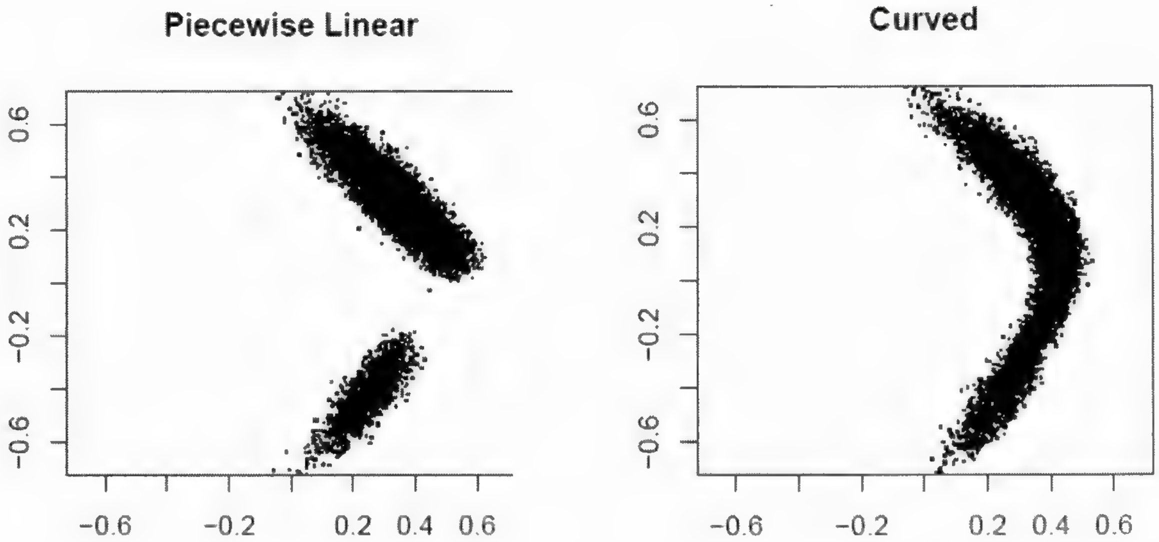
Can be well fit by
e_v, e_w and the derivative.

The sign of the eigenvalues
depends on
the geometry of the
threshold.

Nontrivial thresholds can produce some RGC-type behaviors



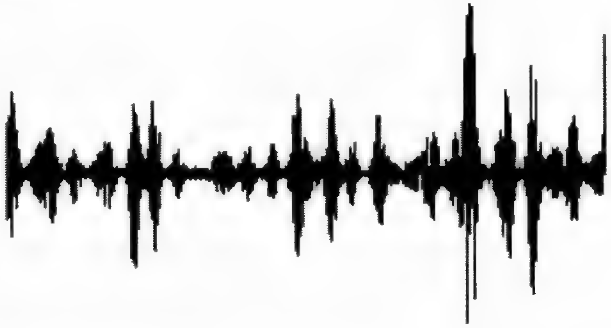
Two negative eigenvalues



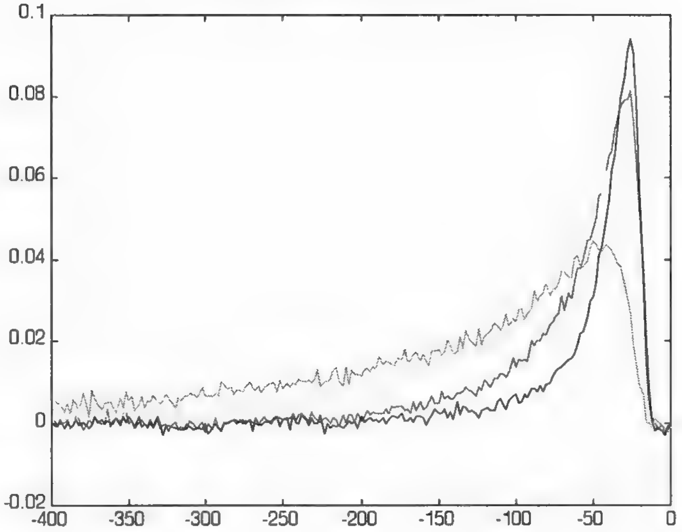
One positive, one negative

Adaptive neural coding

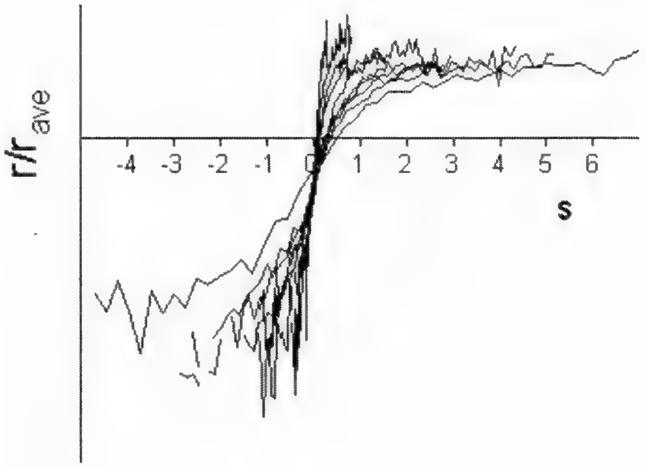
Stimulus s(t)



Features adapt



Threshold functions adapt



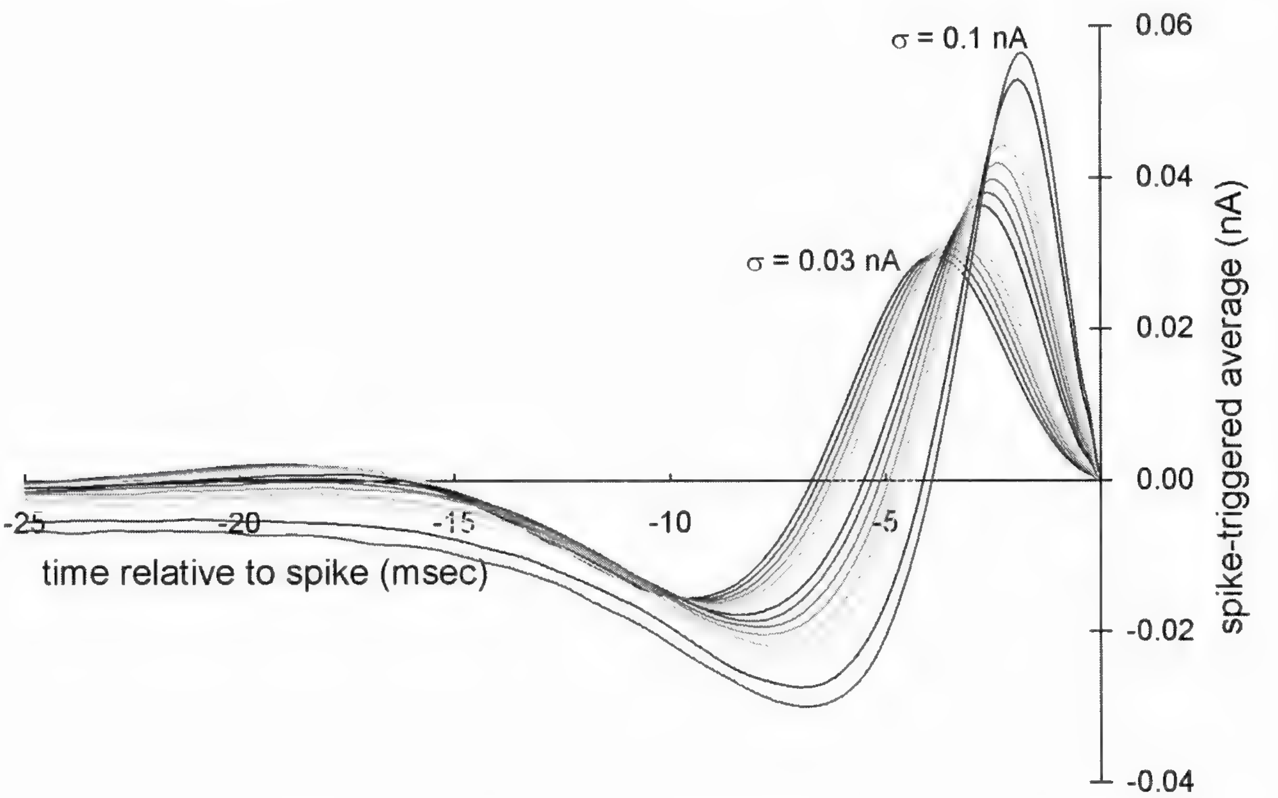
Bill Bialek and Rob de Ruyter van Steveninck

What are potential mechanisms underlying this rapid gain control?

Contribution of intrinsic properties

Yu and Lee, 2000
Kim and Rieke, 2001
Borst and Sompolinsky, 2006
Hong, Aguera y Arcas and Fairhall, 2007
Gaudry and Reinagel, 2007
Diaz and Maravall, 2008

Instantaneous adaptation in the Hodgkin-Huxley model

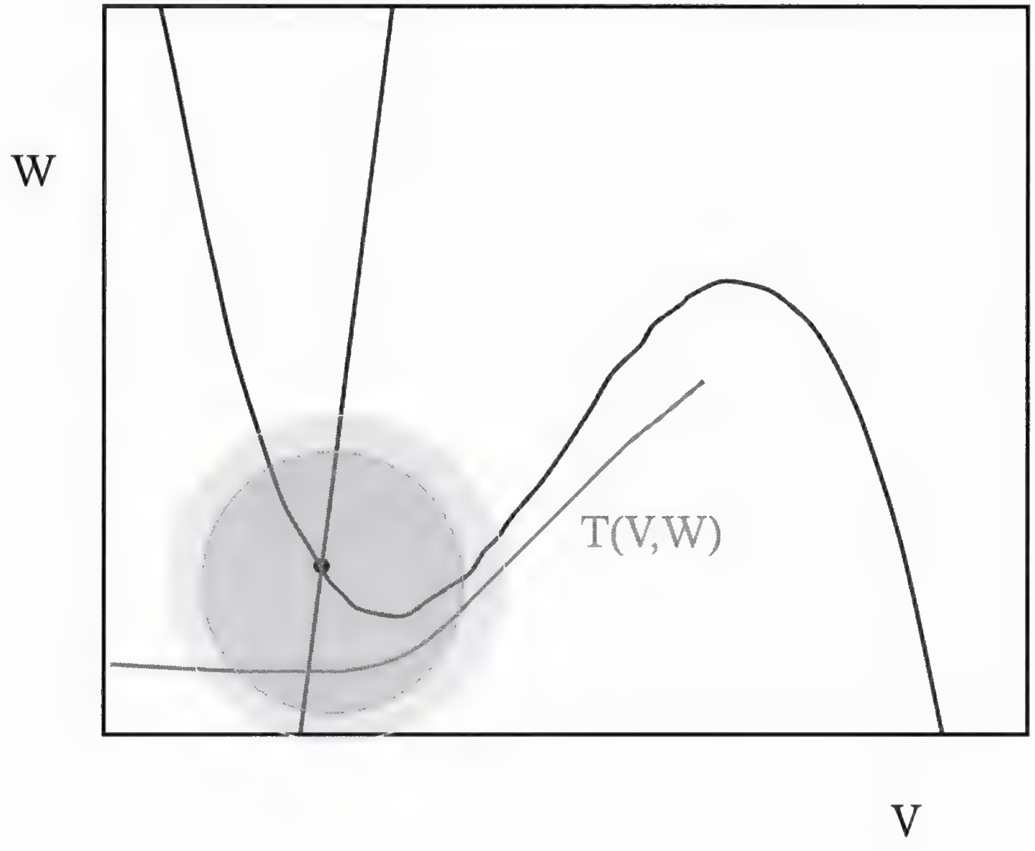


The "relevant feature" of HH adapts to the stimulus variance.

How much adaptation is due to intrinsic neural nonlinearity—
in particular, the curvature of the threshold?

Instantaneous adaptation in the multidimensional LN model

The *m*-LN model allows us to derive explicit variance dependence for the STA and the covariance modes.



Instantaneous adaptation in the multidimensional LN model

The *m*-LN model allows us to derive explicit variance dependence for the STA and the covariance modes.

In this model, changing the variance cannot move you out of the original relevant feature space.

However, the STA will change considerably.

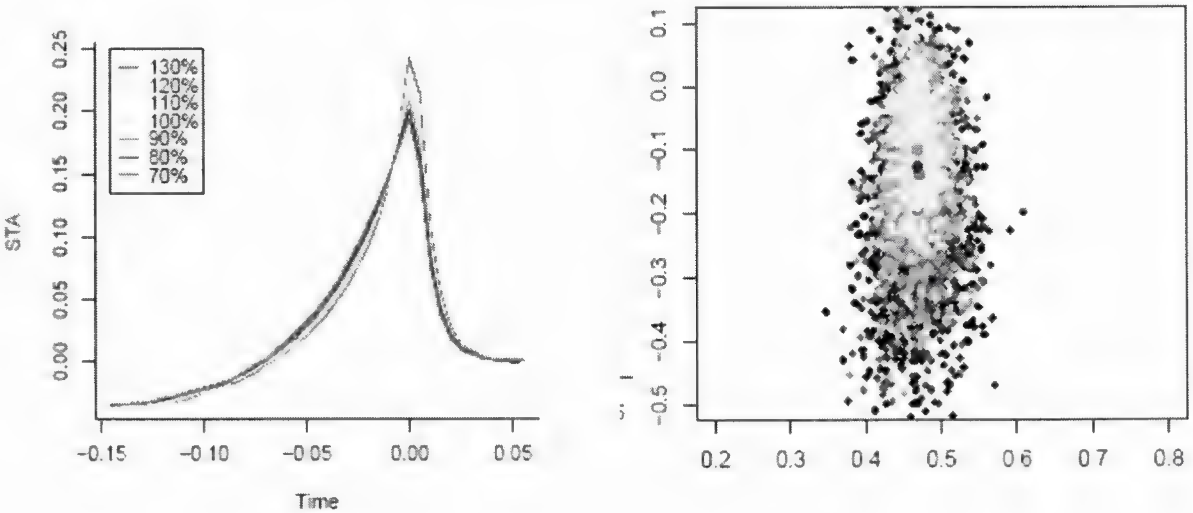
For a *linear* threshold, the derivative component increases with variance.

$$\mathrm{STA}(\tau) = z_0 f_\mu(\tau) + \frac{\sigma}{\sqrt{\pi}} f'_\mu(\tau)$$

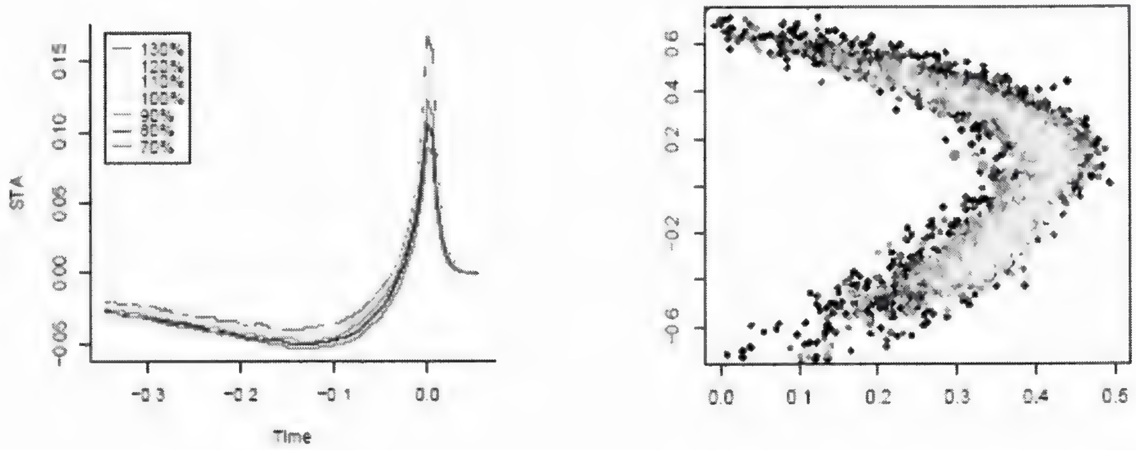
For a *curved* threshold, the centroid of the spike-triggered distribution changes in more complex ways.

Feature adaptation in the multidimensional LN model

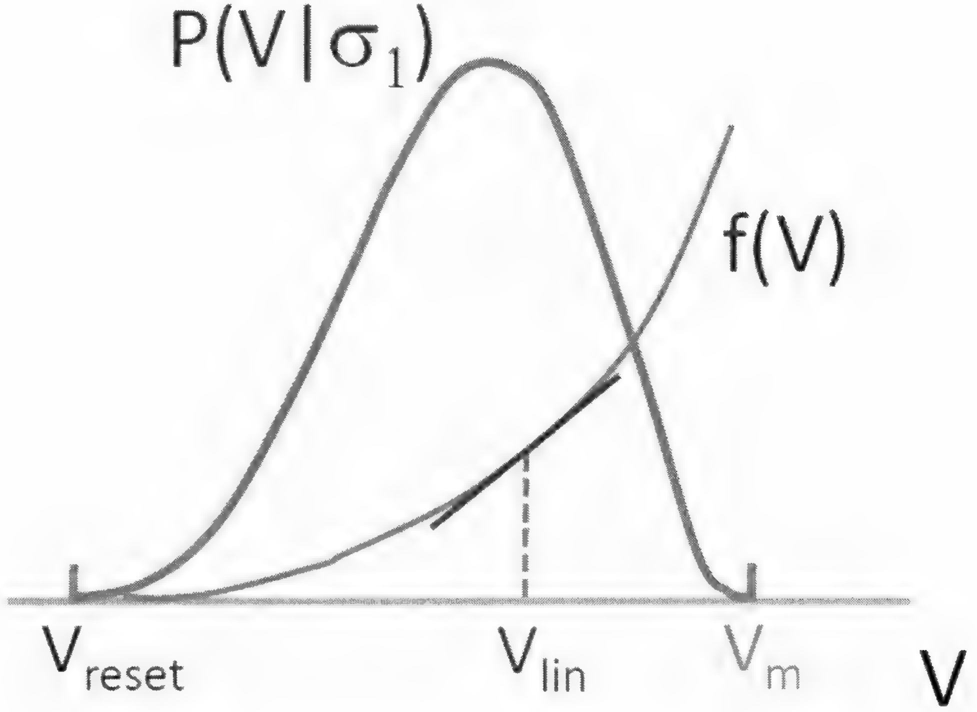
Linear threshold



Curved threshold



Instantaneous adaptation in single neurons



$P(V|\sigma_1)$
f(V)
V_{reset}
V_{lin}
V_m
V

Example: QIF $$\dot{v} = -\frac{1}{\tau_m}v + \alpha v^2 + (v_r - v_s)\,\delta(v - v_s) + \sigma\xi(t)$$

Adaptive spike coding

Single spikes signal the occurrence of particular features in the input

These features and the way they combine nonlinearly to produce spikes are modulated by changes in stimulus statistics

For single neurons, we can derive analytically the form of the relevant features from underlying biophysics and quasi-analytically predict dependence on stimulus statistics

Adaptive coding from neuronal dynamics

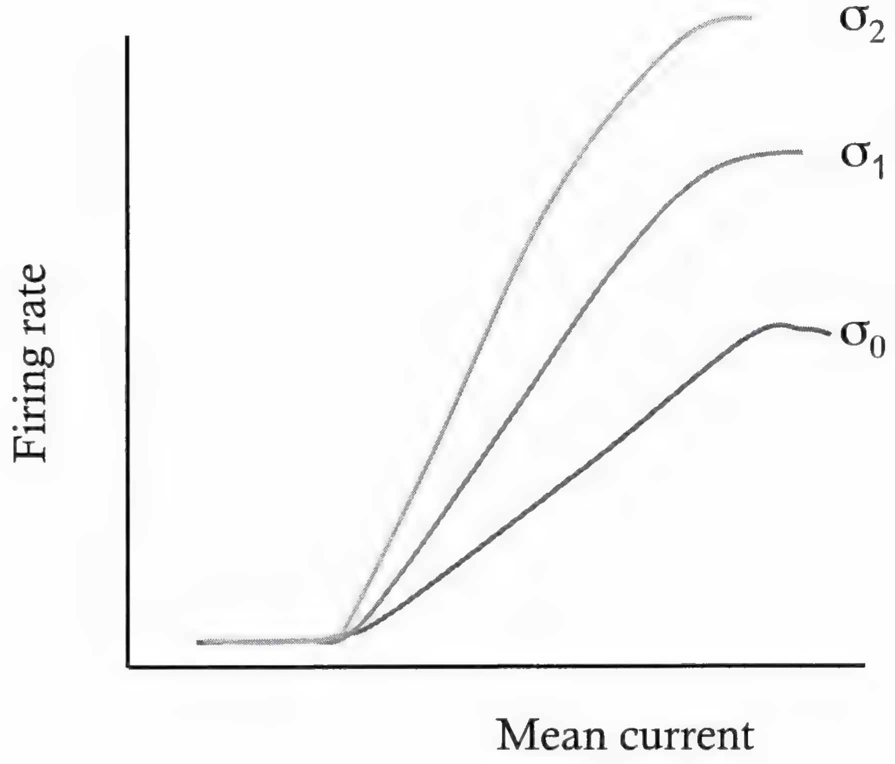
White noise analysis generates functional models of single neural computation

Have shown a limit where the neural dynamical system may be
reduced to a LN model;
    features derive directly from parameters of the system;
    threshold is a consequence of the system's nonlinearity.

As in empirical studies, think of the threshold in multiple dimensions;
    rather than as time-varying.
    Equivalent to a function of V and dV/dt

Showed explicitly how such a model can change its apparent computation as stimulus variance changes.

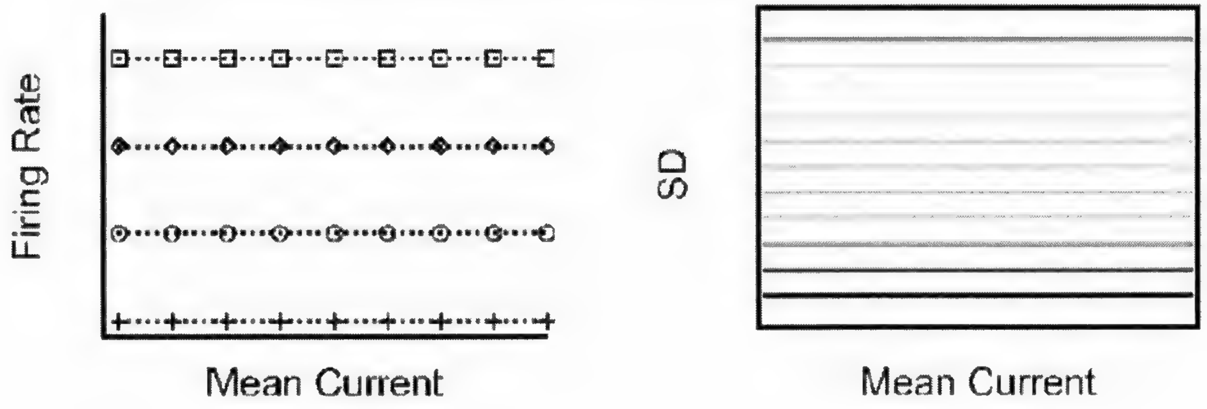
Another type of gain control: steady state rate coding



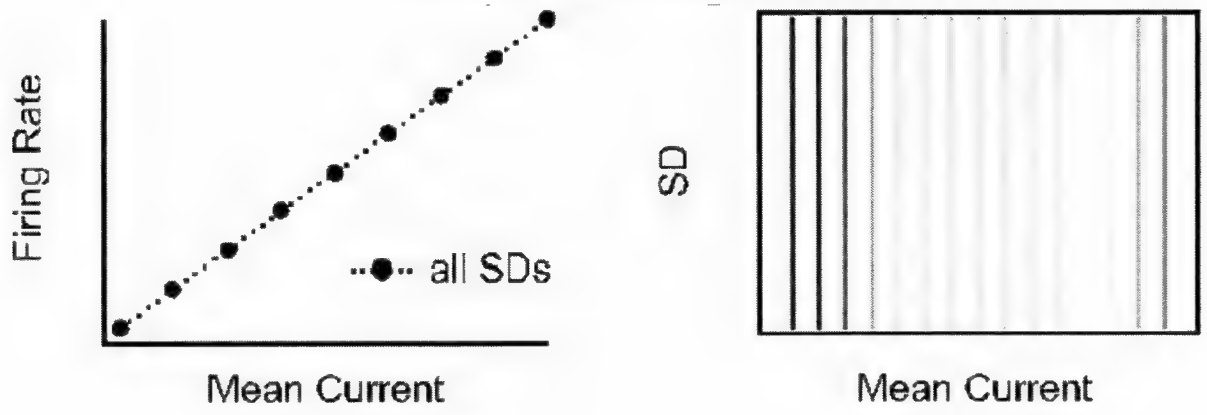
Hong, Lundstrom, Spain, Higgs

Encoding of stimulus statistics: mean vs variance

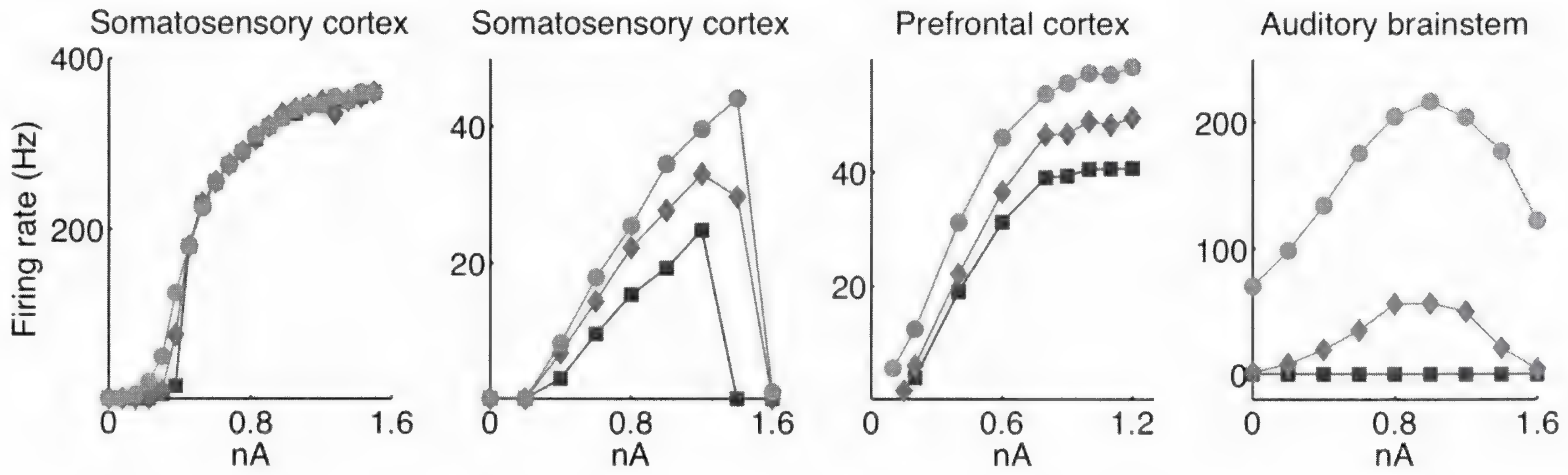
Differentiator or coincidence detector



Temporal integrator

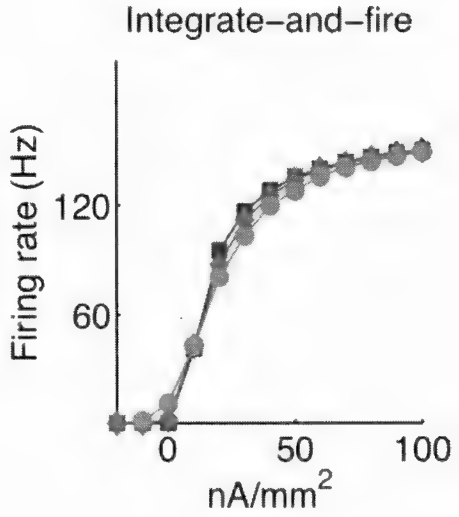


Rate coding in single neurons

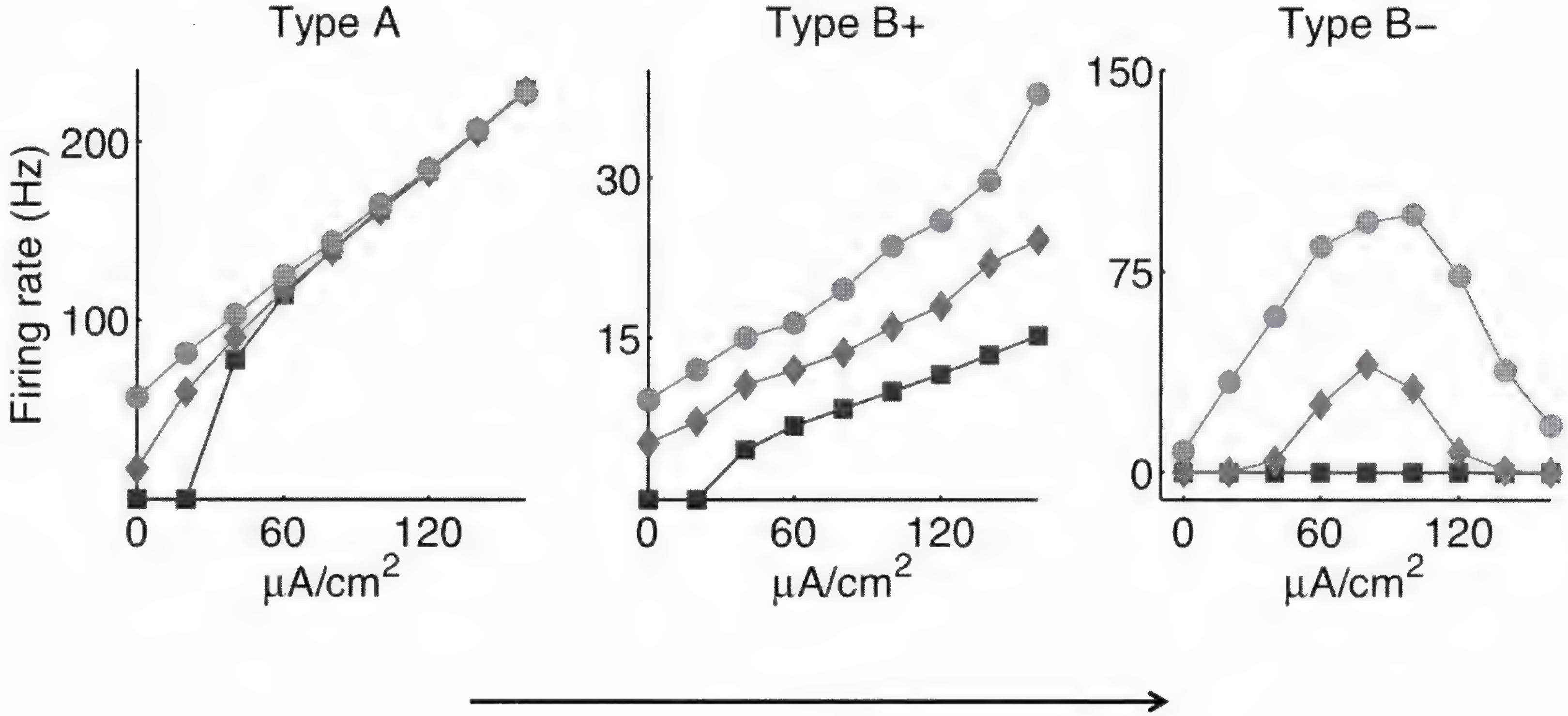


Chance, Abbott and Reyes ‘02; Higgs and Spain ‘06; Arsiero et al. ‘07

How well is this captured by standard model neurons?



Introduce a new classification



HH with a single changing parameter

Work with a reduced model

- Basic framework

$$I_{\text{Total Current}} = I_{\text{Na}} + I_{\text{K}} + I_{\text{Leak}} + I_{\text{Input}}$$

- 4D Hodgkin-Huxley model

$$C\frac{dV}{dt} = -G_{Na}m^3h\left(V - E_{Na}\right) - G_K n^4\left(V - E_K\right) - G_{Leak}\left(V - E_{Leak}\right) + I_{ext}$$

$$\frac{dn}{dt} = \alpha_n(1-n) - \beta_n n$$

$$\frac{dm}{dt} = \alpha_m(1-m) - \beta_m m$$

$$\frac{dh}{dt} = \alpha_h(1-h) - \beta_h h$$

Work with a reduced model

- 2D Hodgkin-Huxley

$$C\frac{dV}{dt} = -G_{\mathrm{Na}} m_{\infty}{}^{3} h\left(V - E_{\mathrm{Na}}\right) - G_{\mathrm{K}} n^{4}\left(V - E_{\mathrm{K}}\right) - G_{\mathrm{Leak}}\left(V - E_{\mathrm{Leak}}\right) + I$$

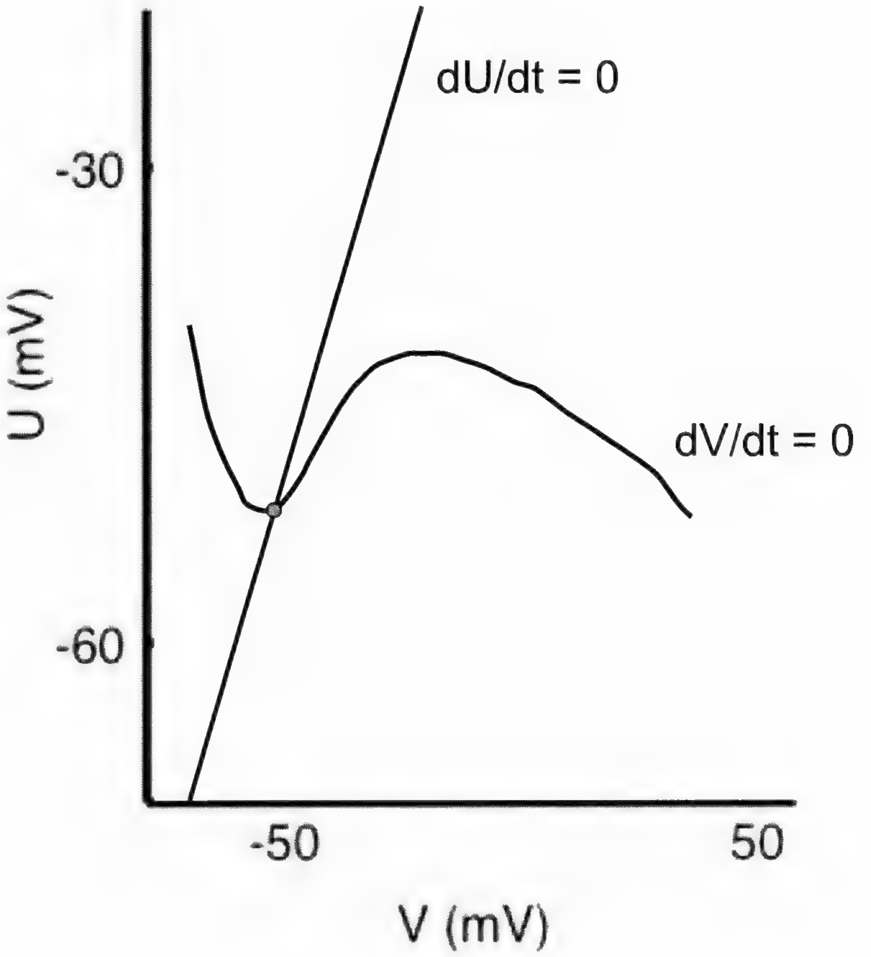
$$\tau\frac{dn}{dt} = n_{\infty} - n$$

$$m_{\infty} = 1/\left(1 + \exp\left((-40 - V)/10\right)\right)$$

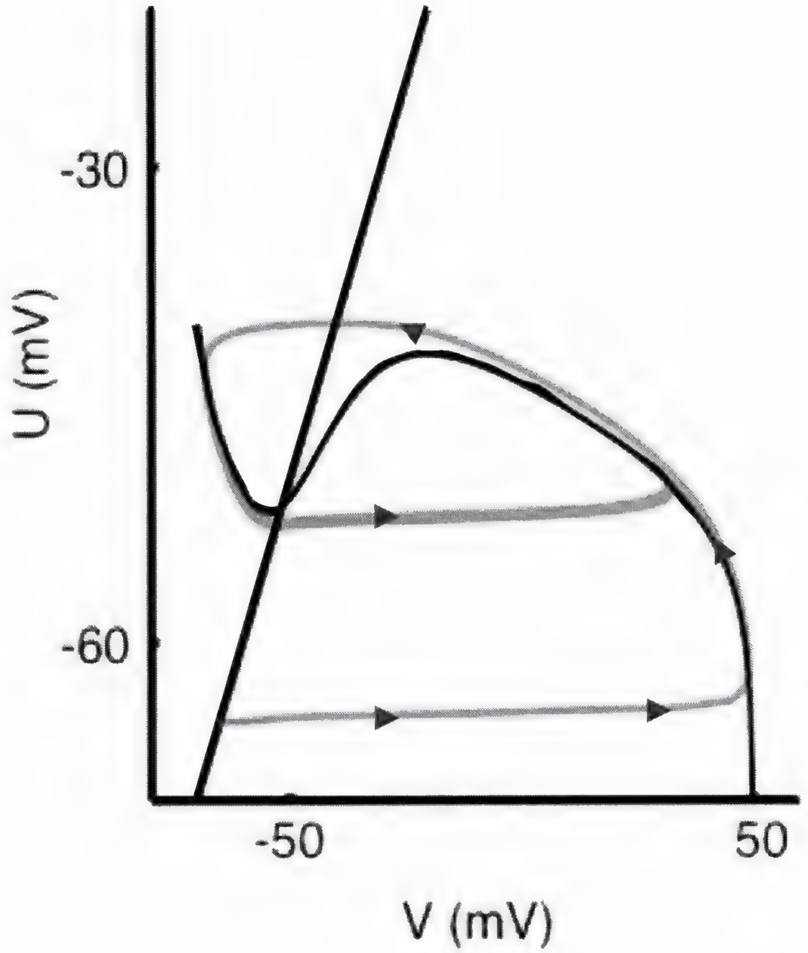
$$h = 0.89 - 1.1n$$

$$n_{\infty} = 1/\left(1 + \exp\left((-50 - V)/15\right)\right)$$

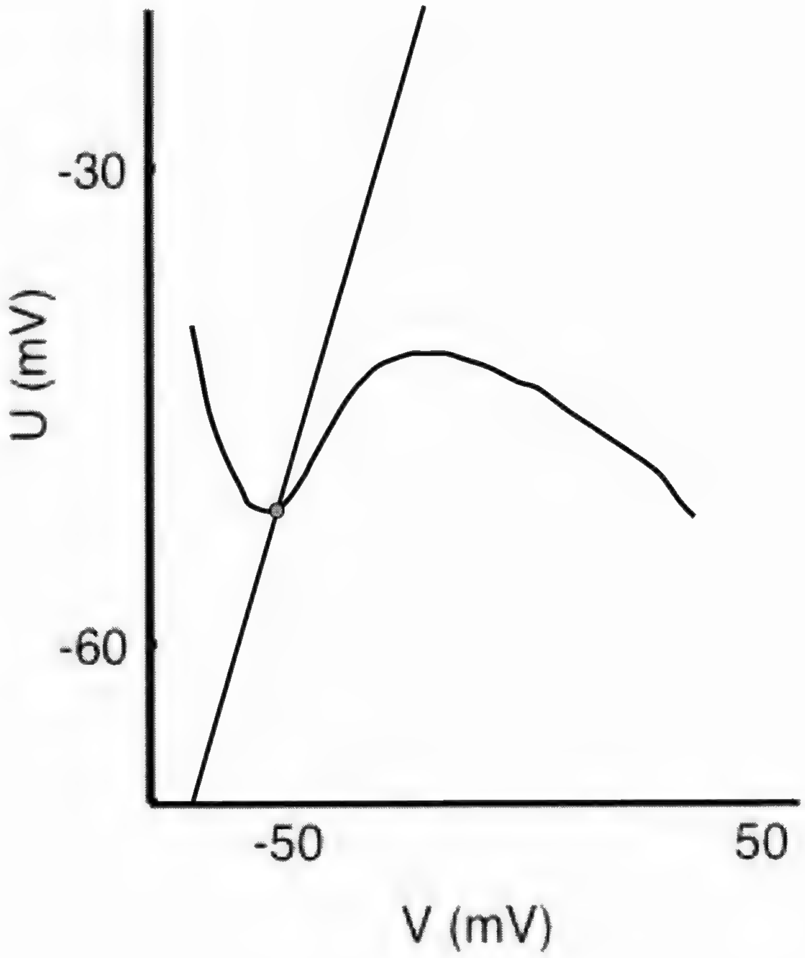
Spiking on the phase plane



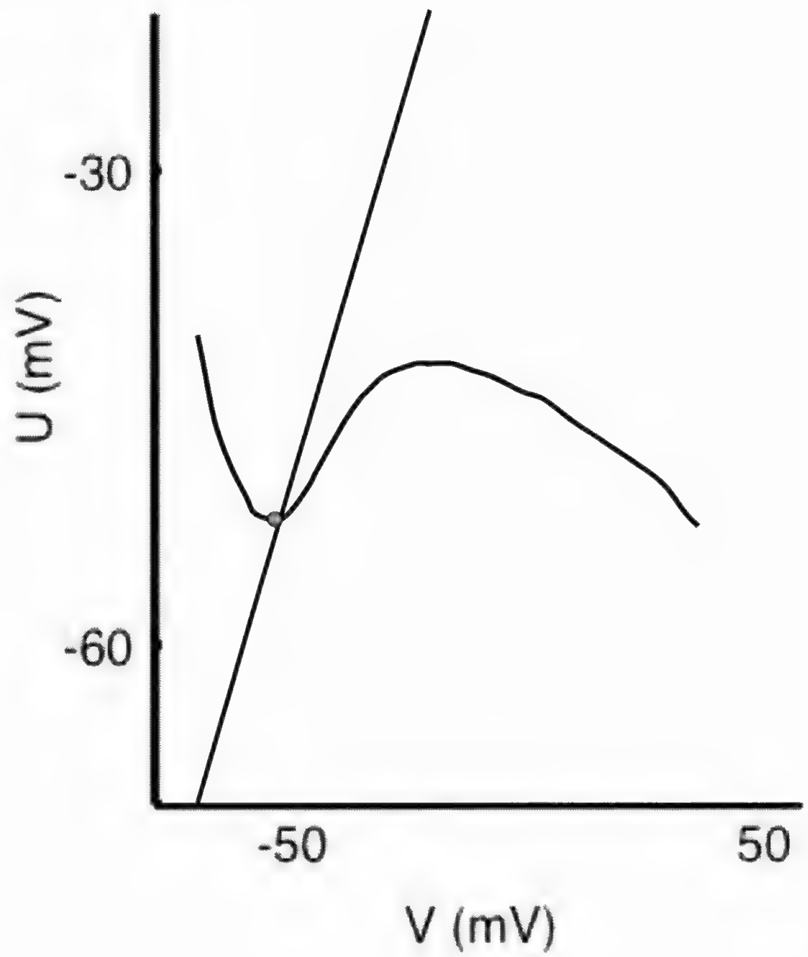
Spiking on the phase plane



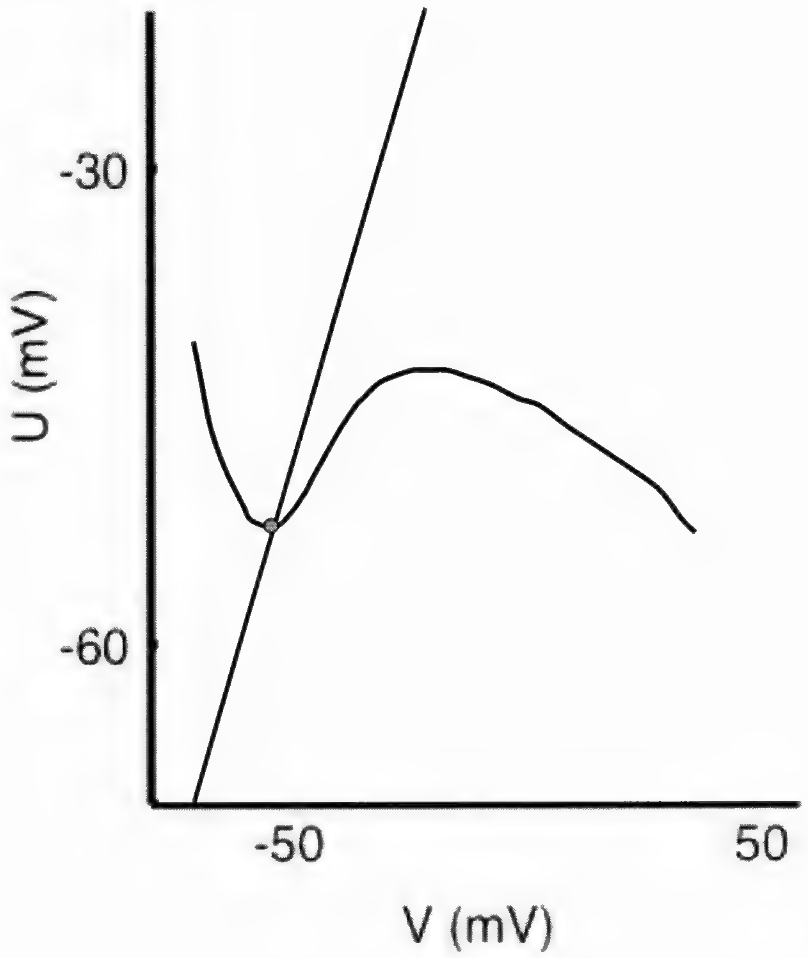
Spiking on the phase plane



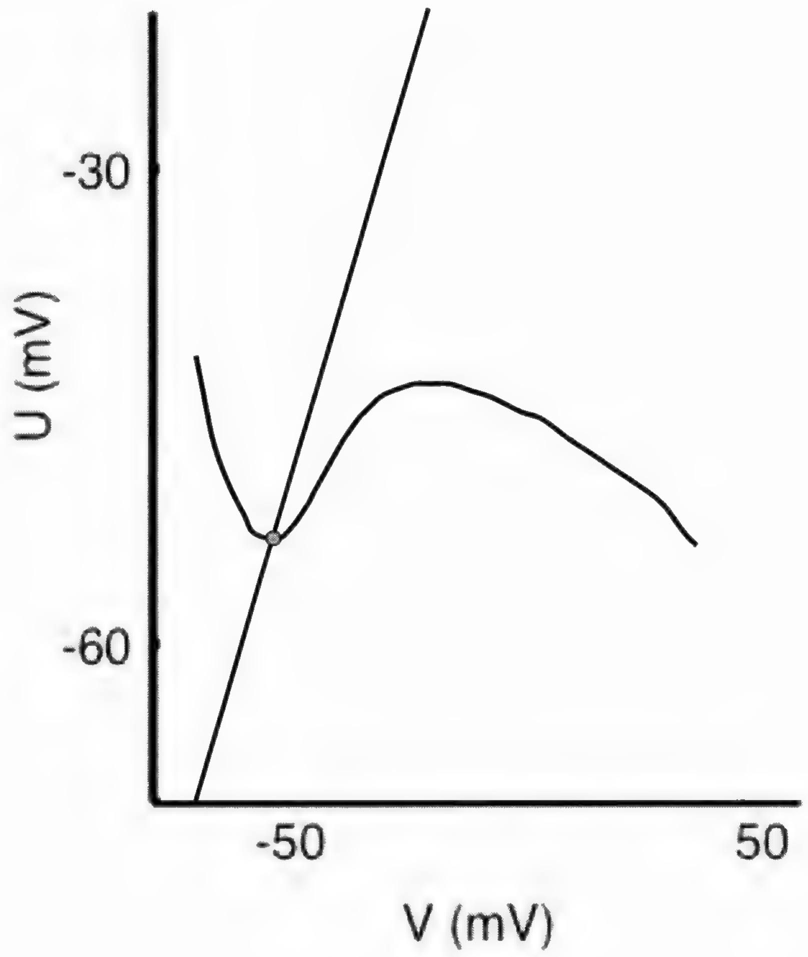
Spiking on the phase plane



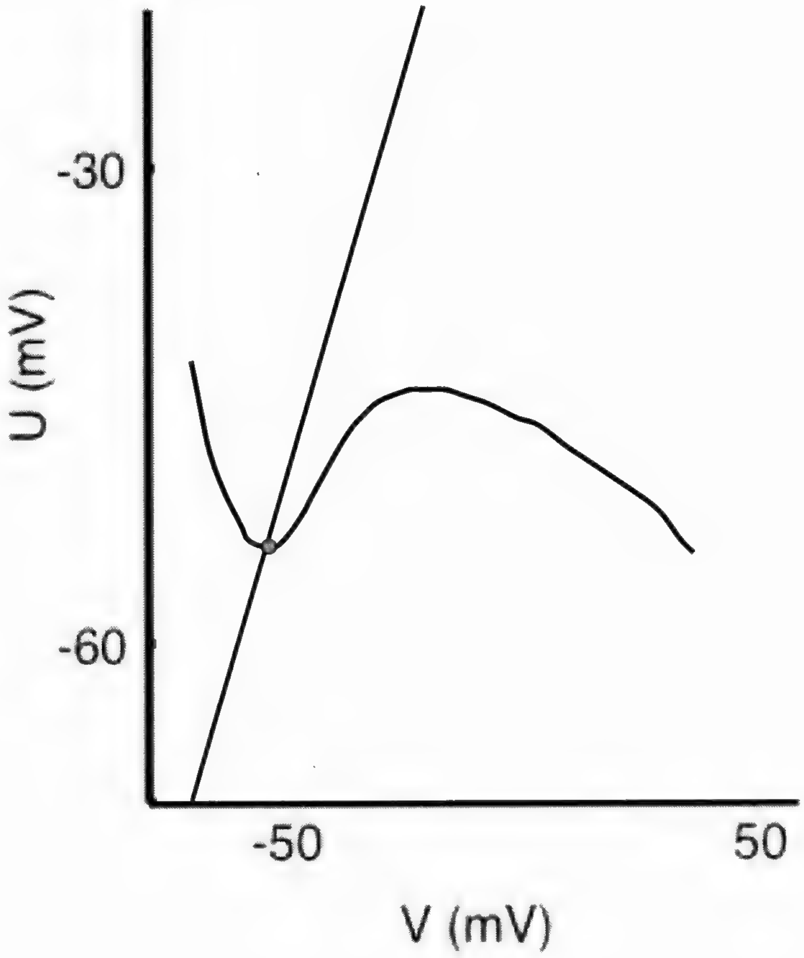
Spiking on the phase plane



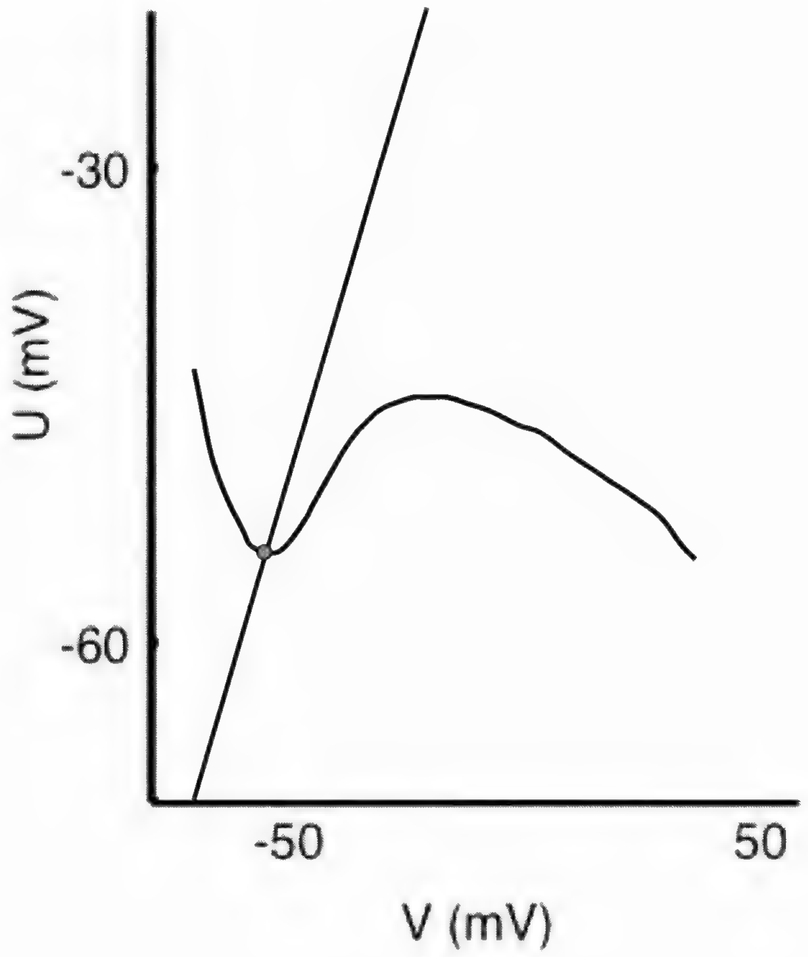
Spiking on the phase plane



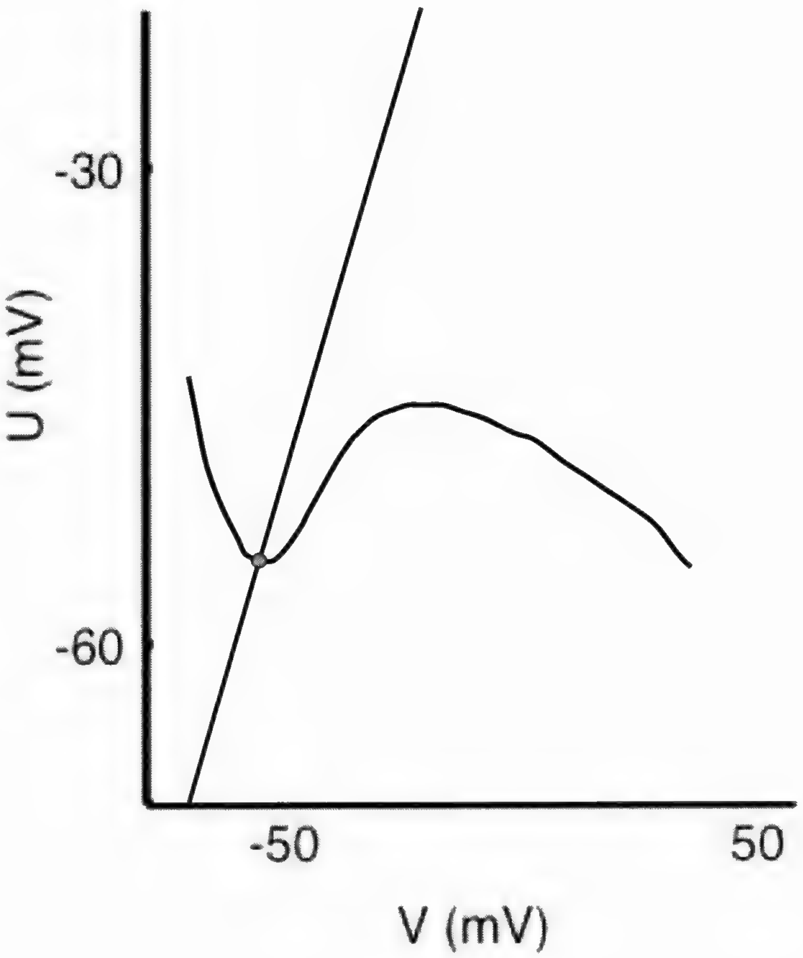
Spiking on the phase plane



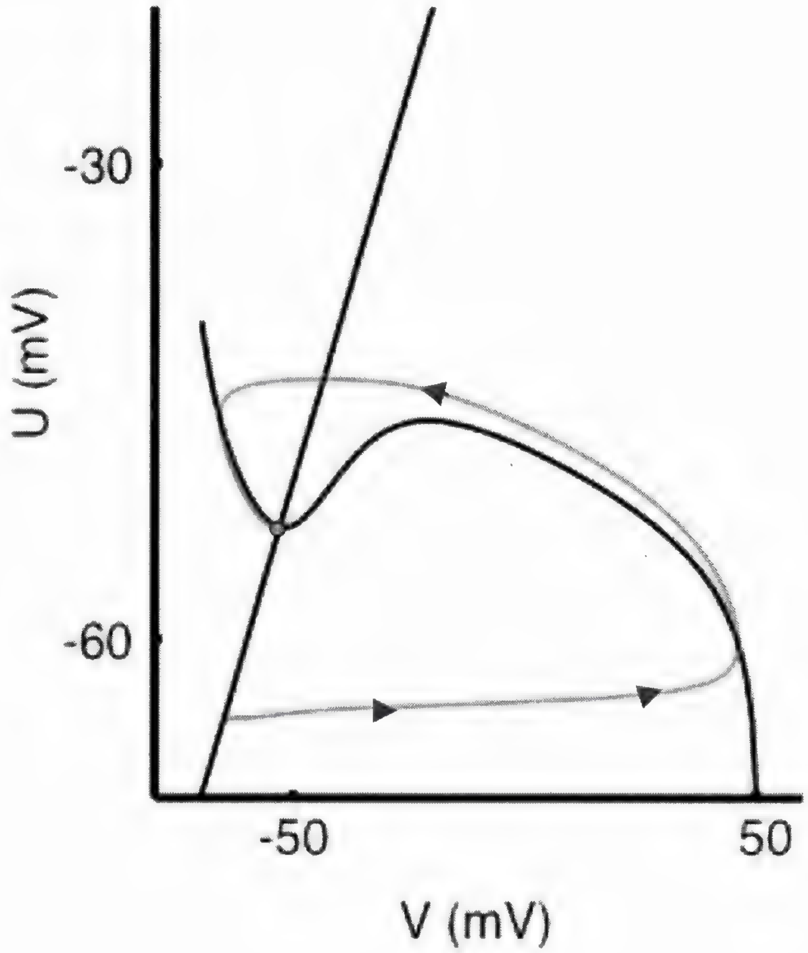
Spiking on the phase plane



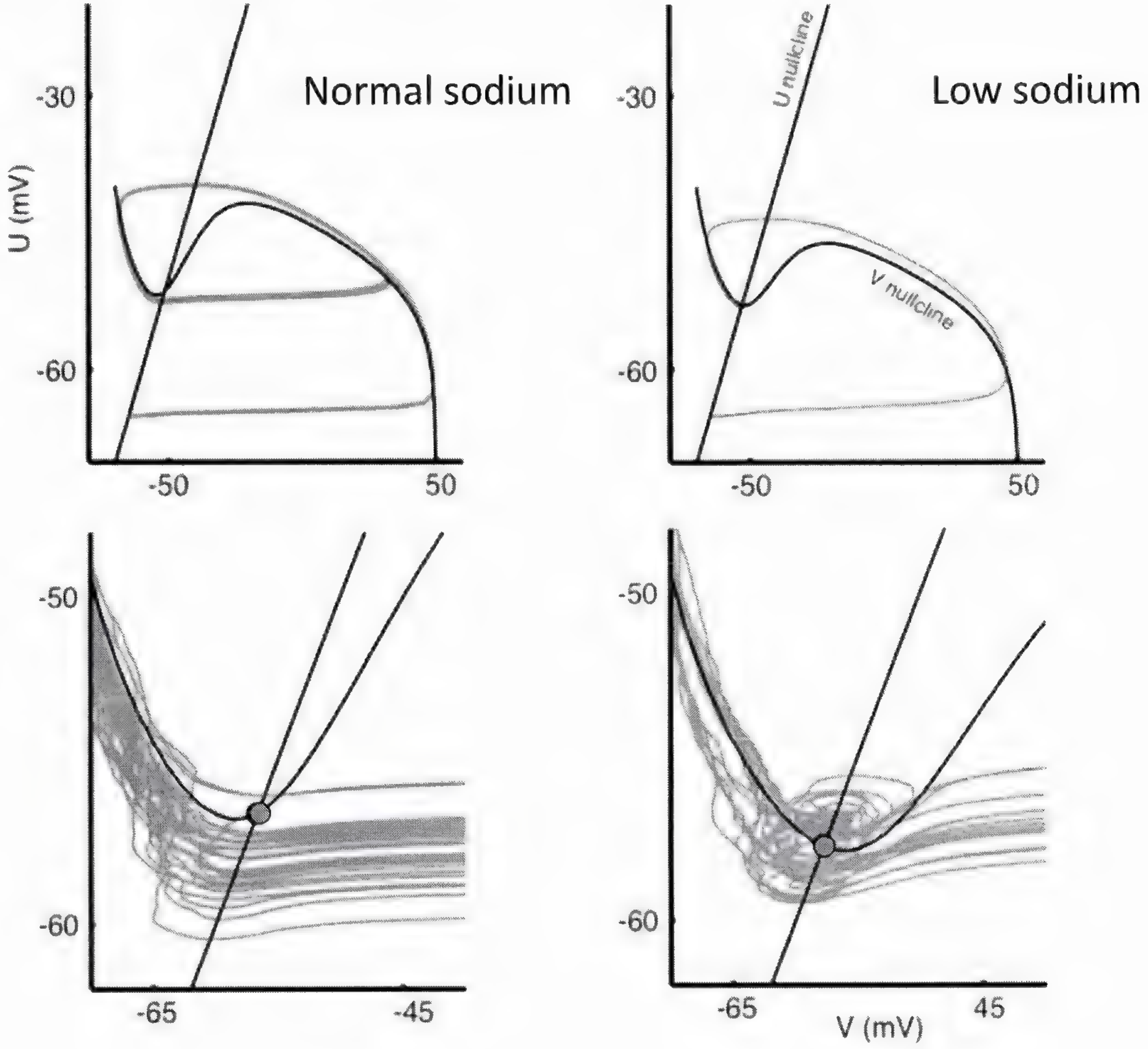
Spiking on the phase plane



Spiking on the phase plane



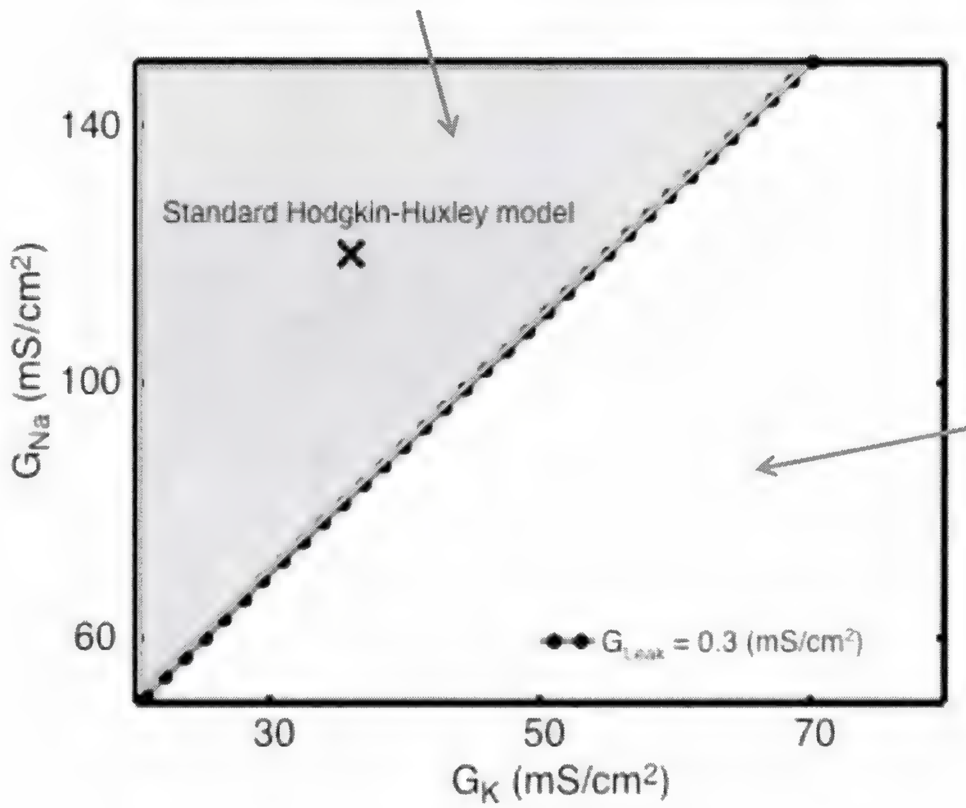
Two types of firing modes in a model neuron



Normal sodium

Low sodium

Boundary of *purely* noise-driven firing in a model neuron

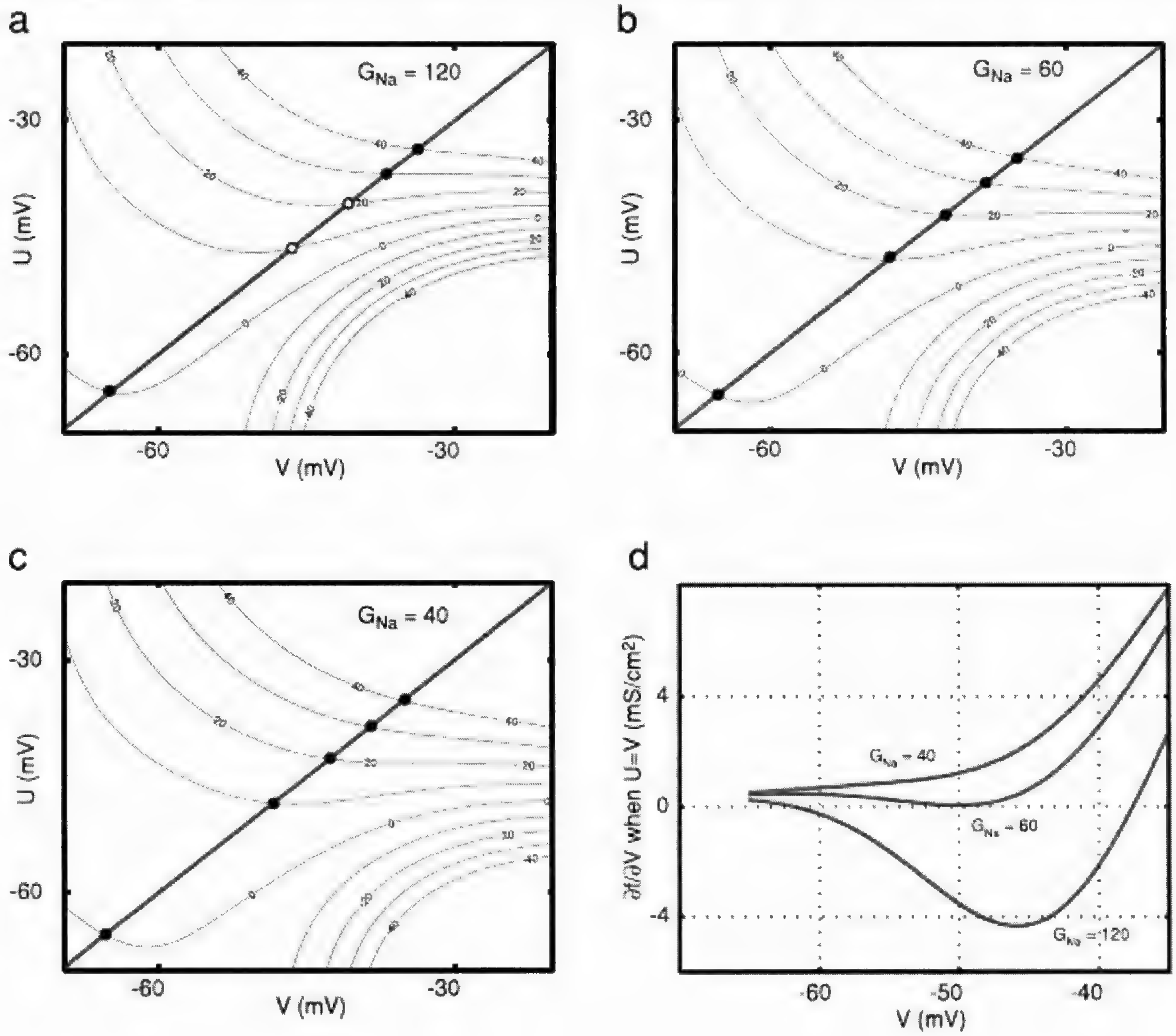


Type A and B+

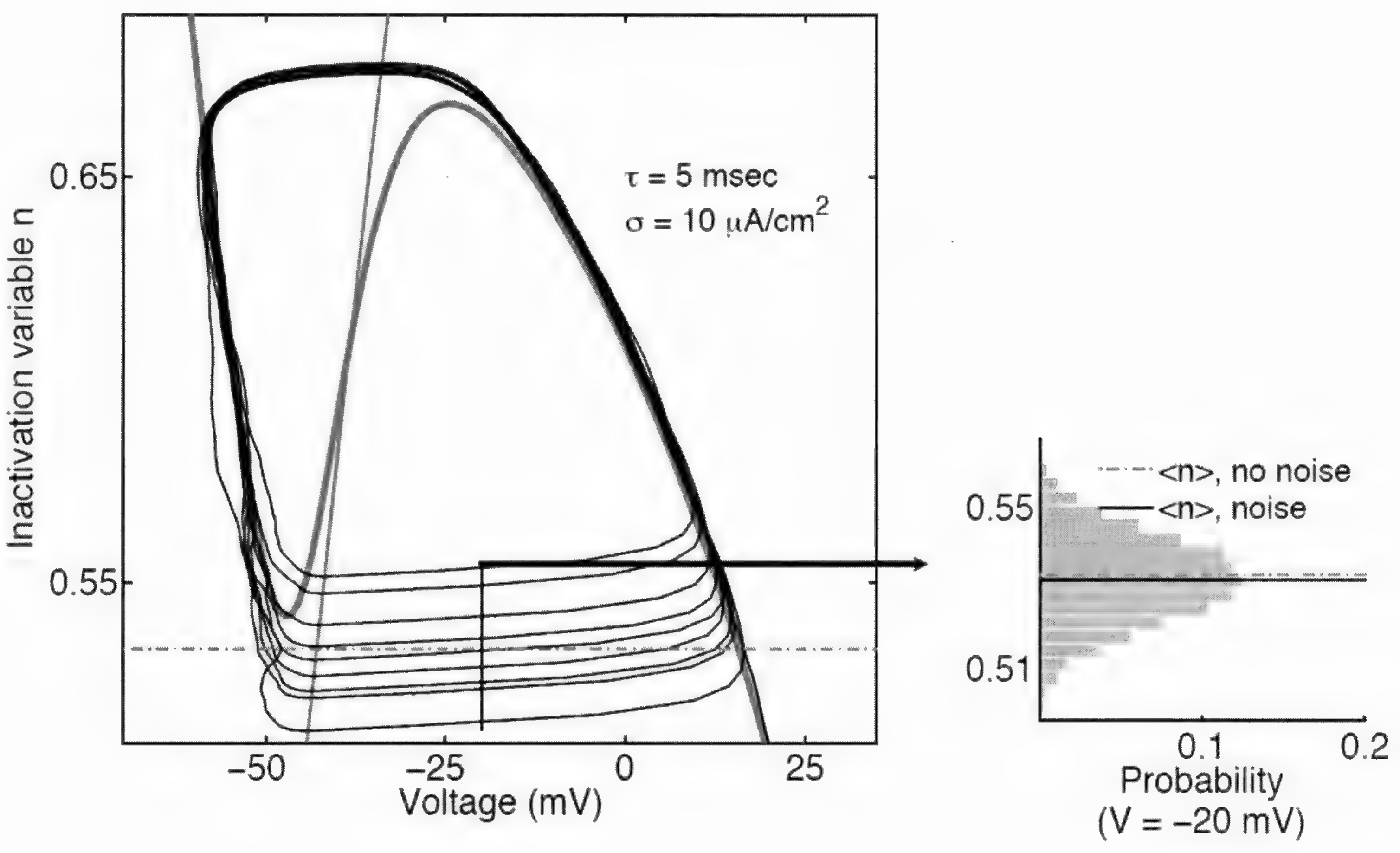
Type B-

Lundstrom, Hong, Higgs and Fairhall, Neural Comp. ‘08

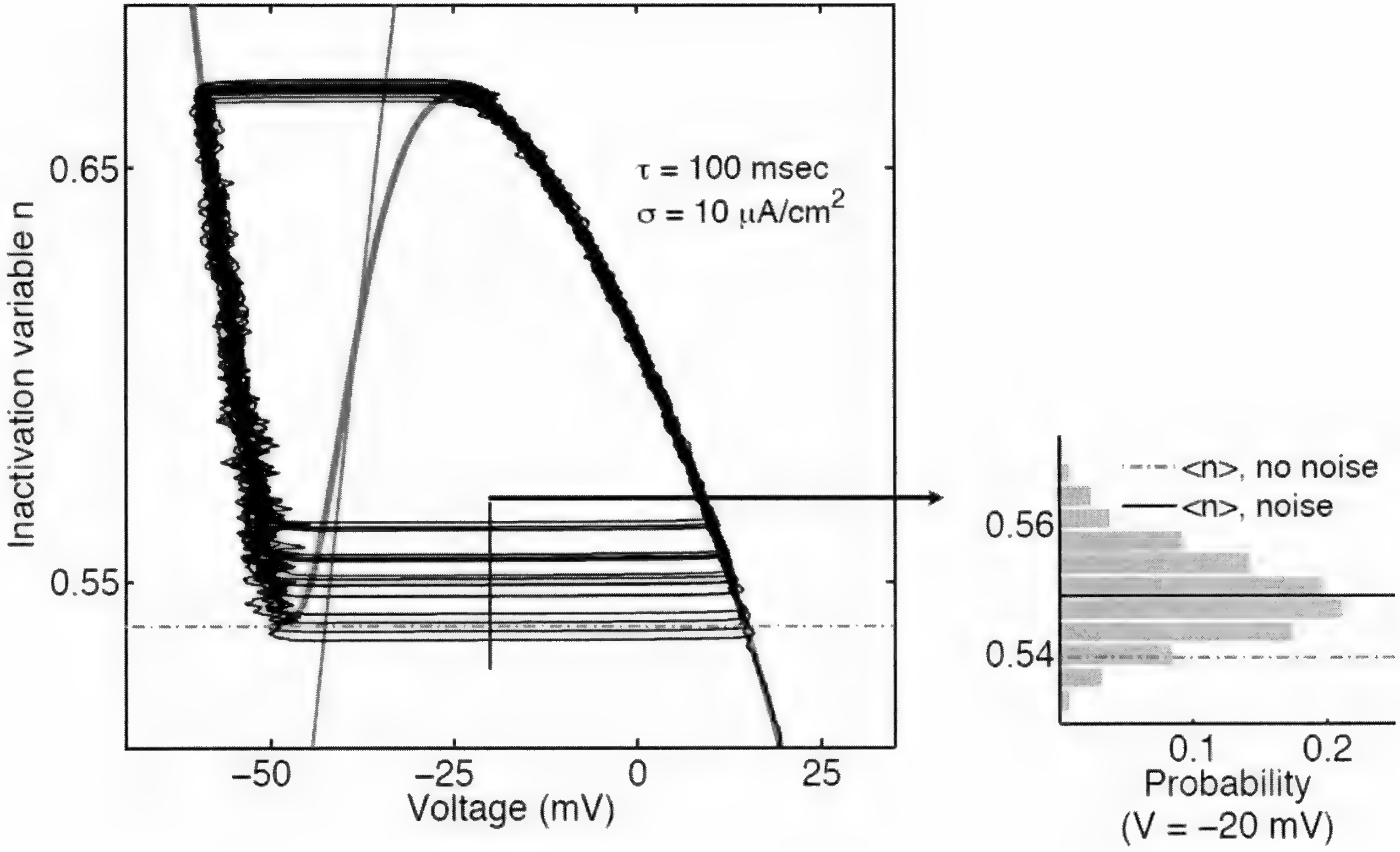
Boundary of *purely* noise-driven firing in a model neuron



What separates noise-sensitive from noise-insensitive?

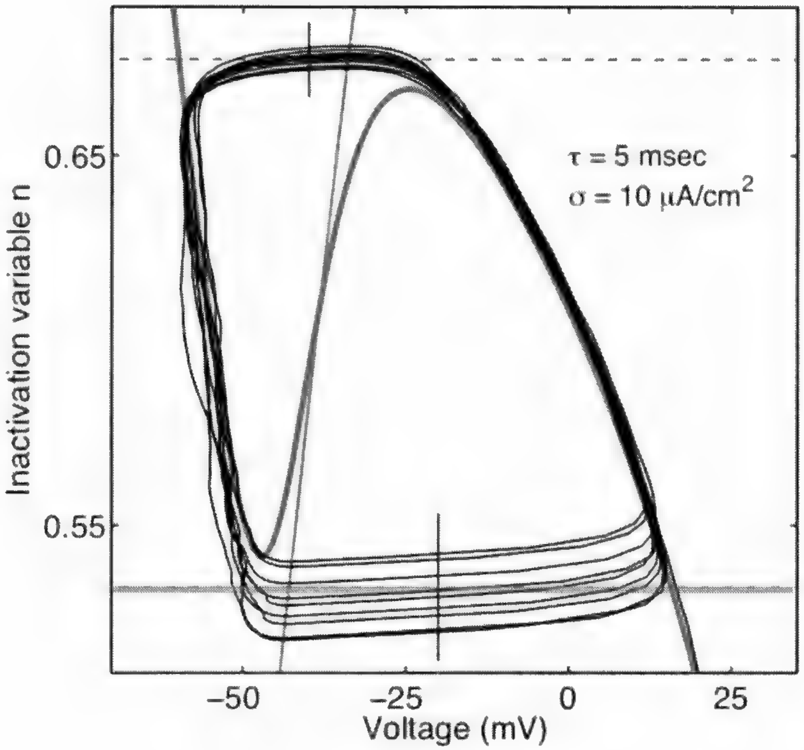


What separates noise-sensitive from noise-insensitive?

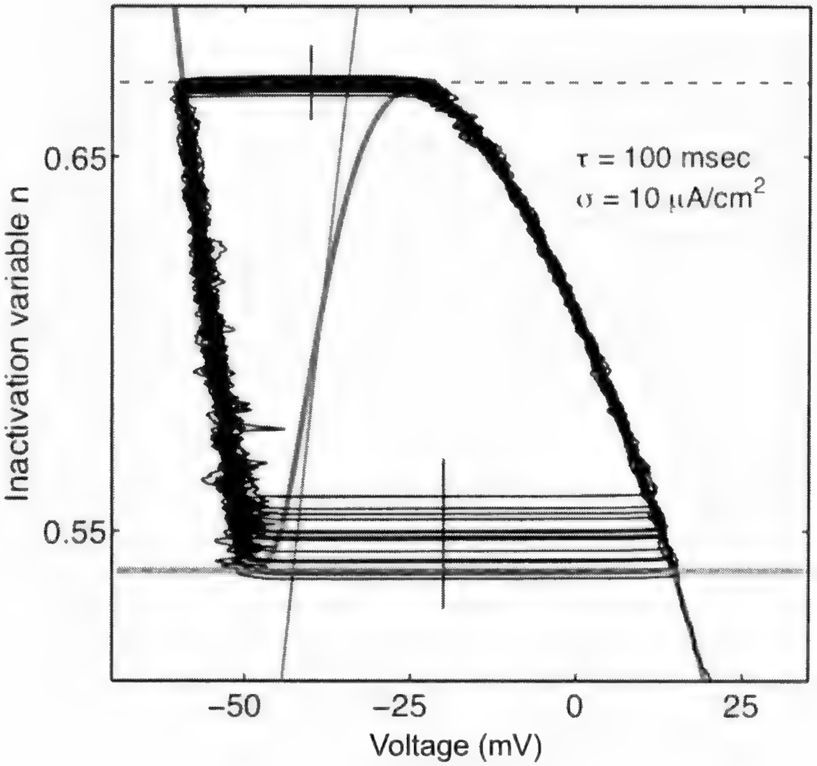


Type A and Type B are separated by a timescale

Type A



Type B



Lundstrom, Famulare, Sorenson and Fairhall, submitted

An even more reduced model

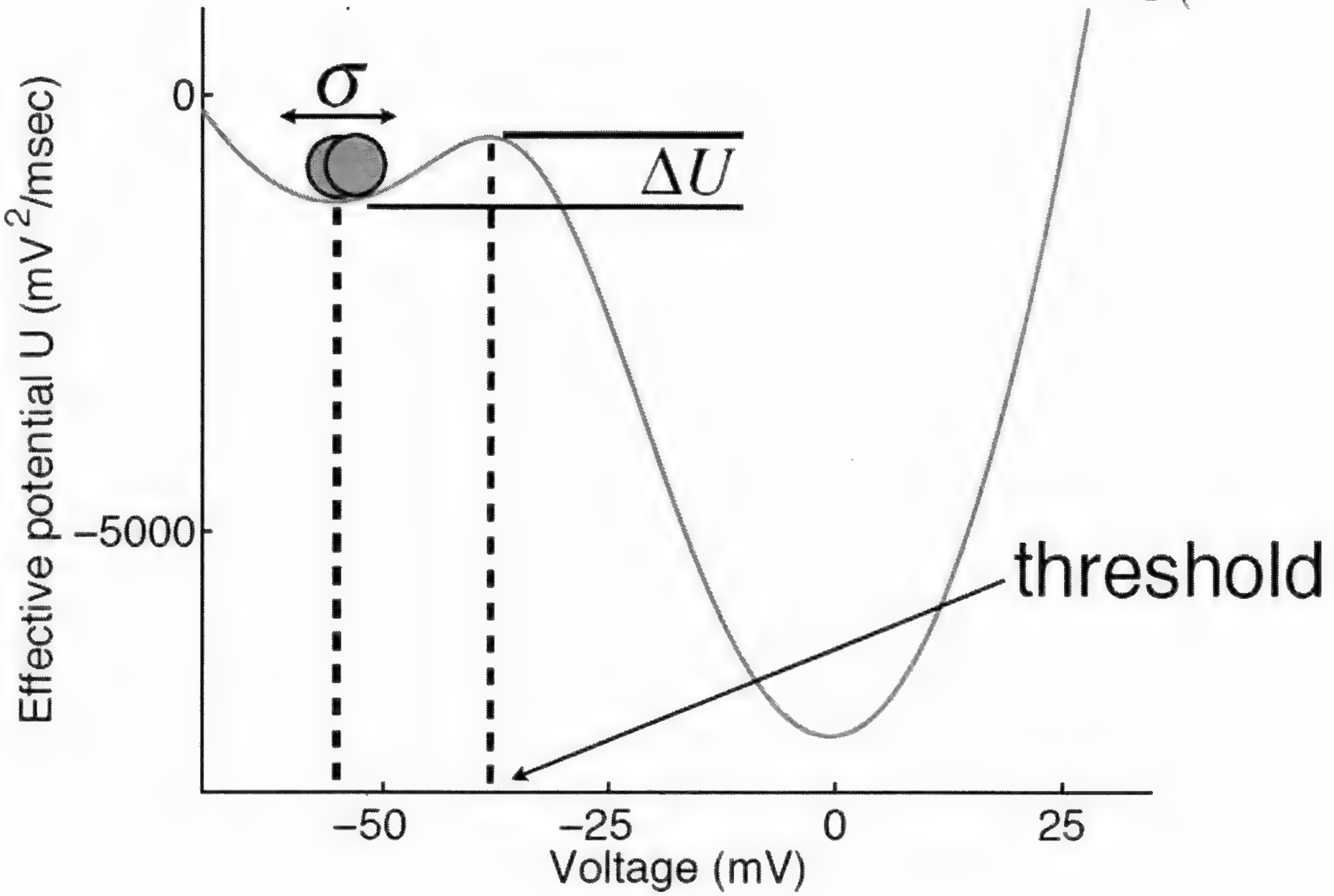
Think about the system, for fixed n, as having a force, $f(V,n) = -dV/dt$

Define the *potential* U such that

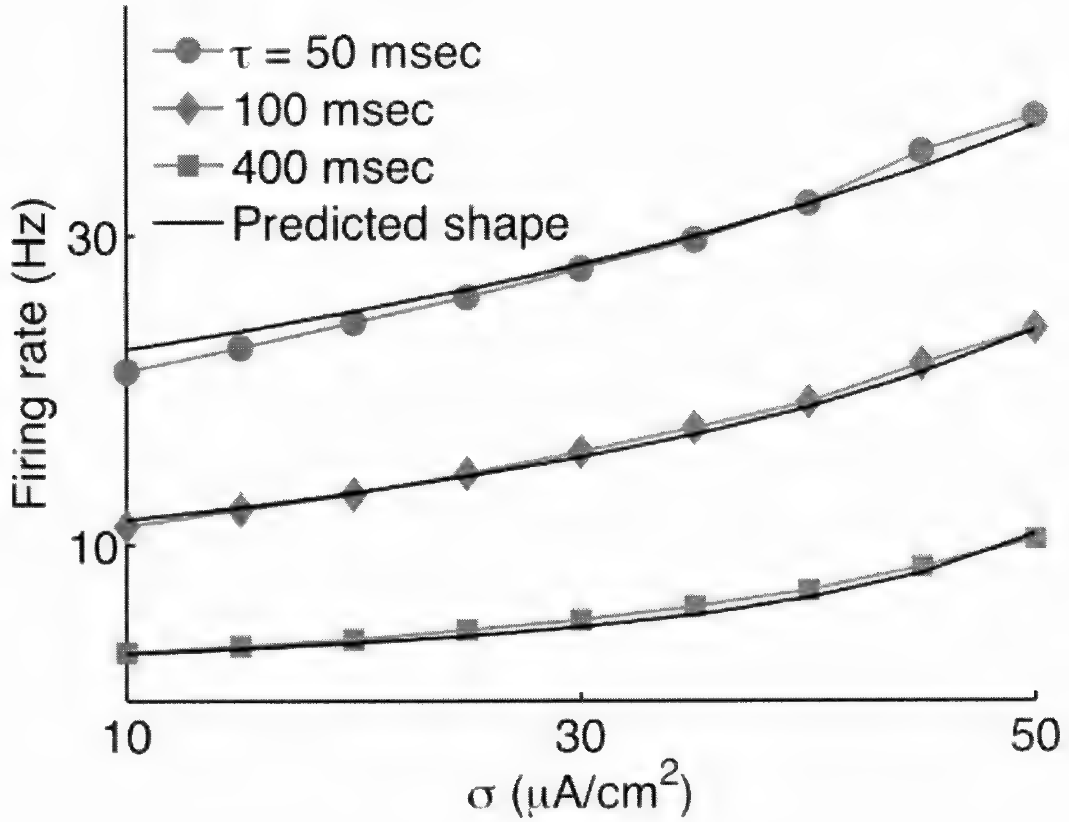
$$\frac{\partial U\left(V,n\right)}{\partial V} = -f\left(V,n\right)$$

An even more reduced model

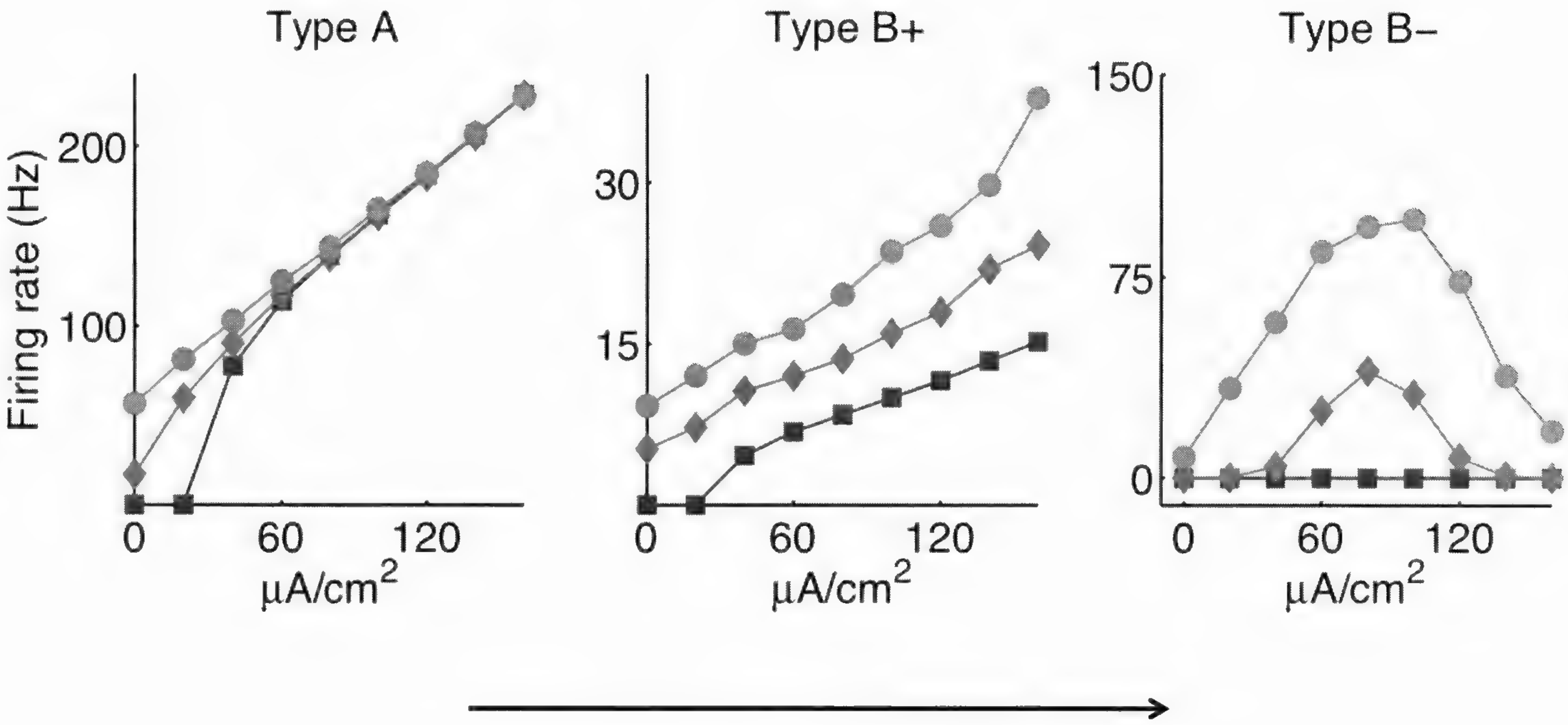
$$1/\tau = k\exp\left(-\Delta U / C\sigma^2\right)$$



Sensitivity to fluctuations is predicted as a function of τ



Back to our classification

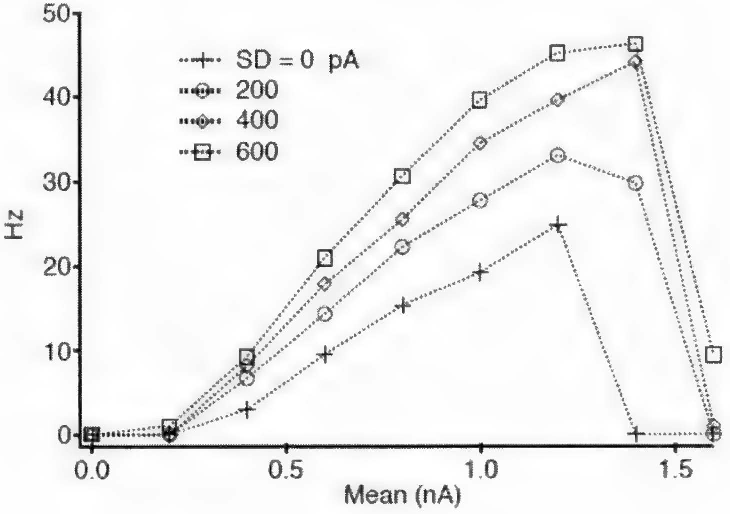


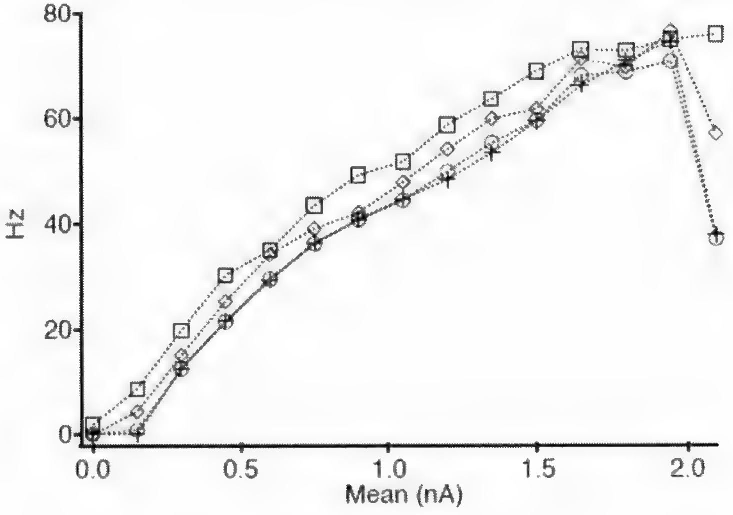
HH with slow sodium inactivation:
changing G_{Na} parameter

Slow channels confer sensitivity to fluctuations

Neurons are more sensitive to input fluctuations when:

- Ratio of K channel currents to Na channel currents is large
- Slower adaptation currents are present
  - Slow after-hyperpolarization (sAHP) current


pyramidal neuron with large sAHP


pyramidal neuron with small sAHP

Higgs *et al.*, J Neurosci, 2006

From dynamics to coding

Different definitions of coding in terms of f-I curves vs LN models

Linking biophysical parameters to coding properties in simple models:

-- transitions between different behaviors

-- determining the structure of encoded features

-- combination of dynamical systems and stochastic analysis allows us to start to explore how statistics and dynamics interact

ARTICLE Communicated by Tatyana Sharpee

Single Neuron Computation: From Dynamical System to Feature Detector

Sungho Hong
shhong@u.washington.edu
Department of Physiology and Biophysics, University of Washington, Seattle, WA 98195, U.S.A.

Blaise Agüera y Arcas
blaisea@microsoft.com
Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544

Adrienne L. Fairhall
fairhall@u.washington.edu
Department of Physiology and Biophysics, University of Washington, Seattle, WA 98195, U.S.A.

White noise methods are a powerful tool for characterizing the computation performed by neural systems. These methods allow one to identify the feature or features that a neural system extracts from a complex input and to determine how these features are combined to drive the system's spiking response. These methods have also been applied to characterize the input-output relations of single neurons driven by synaptic inputs, simulated by direct current injection. To interpret the results of white noise analysis of single neurons, we would like to understand how the obtained feature space of a single neuron maps onto the biophysical properties of the membrane, in particular, the dynamics of ion channels. Here, through analysis of a simple dynamical model neuron, we draw explicit connections between the output of a white noise analysis and the underlying dynamical system. We find that under certain assumptions, the form of the relevant features is well defined by the parameters of the dynamical system. Further, we show that under some conditions, the feature space is spanned by the spike-triggered average and its successive order time derivatives.

1 Introduction

A primary goal of sensory neurophysiology is to understand how stimuli from the natural environment are encoded in the spiking output of

Neural Computation **19, 3133–3172** (2007) © 2007 Massachusetts Institute of Technology

neurons. A useful tool for performing this characterization is white noise analysis (Marmarelis & Marmarelis, 1974; Wiener, 1958; Rieke, Warland, Bialek, & de Ruyter van Steveninck, 1997), whereby a system is stimulated with a randomly varying broadband signal and the relevant features driving the system are determined by correlating output spikes with the signal. Extensions of this analysis to second order have been used to find multiple stimulus features to which the system is sensitive (de Ruyter van Steveninck & Bialek, 1988; Brenner, Bialek, & de Ruyter van Steveninck, 2000; Bialek & de Ruyter van Steveninck, 2005). Recently, these methods have been applied to characterize the computation performed by a single neuron on its current inputs (Agüera y Arcas, Fairhall, & Bialek, 2003; Agüera y Arcas & Fairhall, 2003; Slee, Higgs, Fairhall, & Spain, 2005).

Due to detailed knowledge of the dynamics of ion channels, it is possible to build dynamical models of current flow in neurons that accurately reproduce the voltage response of single neurons to current (or conductance) inputs. White noise analysis provides us with the capacity to reduce this complex set of dynamical equations to a simple functional model with concrete components: linear feature selection followed by a nonlinear decision function generating spikes. This procedure is compelling as it provides a clear intuition about the form of the changes in the stimulus to which the system is sensitive and how the system combines these features in the decision to fire. Second order white noise methods have provided insight into coding properties at the systems level, in visual (Brenner et al., 2000; Fairhall et al., 2006; Touryan, Lau, & Dan, 2002; Rust, Schwartz, Movshon, & Simoncelli, 2005; Horwitz, Chichilnisky, & Albright, 2005) and somatosensory (Petersen, 2003; Maravall, Peterson, Fairhall, Arabzadeh, & Diamond, 2007) cortex. More recently, these methods have also been applied to characterize neural coding in single auditory (Slee et al. 2005), central (Powers, Dai, Bell, Percival, & Binder, 2005), and model (Agüera y Arcas, et al. 2003; Agüera y Arcas & Fairhall, 2003) neurons. However, mostly missing from these analyses is a clear link between the obtained stimulus features and the biophysical properties of the circuit or neuron. This issue is particularly well posed for single neurons under stimulation by direct somatic current injection, where the biophysical parameters of the soma must dominate the neuron's computation. Thus, given the power of white noise methods, the questions that we wish to address here are: How does the dynamical system governing neural behavior map to the neuron's functional characterization derived using white noise analysis? How are the features determined by the neuron's biophysical properties?

1.1 Minimal Spiking Models. Neural dynamics are typically described by conductance-based models that describe the temporal evolution of the voltage V due to an input current I and the variable ionic conductances of ion channels:

$$CdV/dt = -\sum_i g_i(V)(V - E_i) + I, \tag{1.1}$$

where C is the membrane capacitance. Each conductance type i has a reversal potential E_i, where the conductances g_i corresponding to different ion channels may be voltage dependent through the dynamics of the corresponding activation and inactivation gating variables ξ_i and ϕ_i:

$$g_i = \bar{g}_i \xi_i^{p_i} \phi_i^{q_i} \tag{1.2}$$

$$d\xi_i/dt = f_{\xi_i}(\xi, V) \tag{1.3}$$

$$d\phi_i/dt = f_{\phi_i}(\phi, V), \tag{1.4}$$

where $\bar{g}_i$ are constants and the functions f_ξ and f_ϕ, are affine in ξ and ϕ, respectively, but not in V. The exponents p_i and q_i are usually taken to be integers and describe the cooperativity of molecules required to gate the ion channel. This set of equations is highly nonlinear due to this cooperativity and the voltage dependence of the conductances. While ultimately we would like to understand systems of many dynamical variables, we will begin by analyzing a simplified spiking system: the FitzHugh-Nagumo (FN) neural model (FitzHugh, 1961; Nagumo, Arimoto, & Yoshikawa, 1962). The FitzHugh-Nagumo system is cubic in V, the lowest-order nonlinearity supporting spike-like behavior, and is defined by

$$\psi\, dV/dt = V(1 - V)(a + V) - W + c + I, \tag{1.5}$$

$$dW/dt = V - bW, \tag{1.6}$$

where V is voltage, W is a generic inactivation variable with linear recovery dynamics, I is the input current, and a, b, c, and ψ are parameters, $\psi \ll 1$. The nullclines of the system are the curves along which $dV/dt = 0$ (the V nullcline, a cubic) and $dW/dt = 0$ (the W nullcline, a straight line). Figure 1a shows the nullclines for zero input I on the phase plane. Trajectories have been traced forward and backward in time through a randomly chosen set of initial conditions. The unique intersection of the nullclines is the stable fixed point. Elsewhere in the plane, trajectories always spiral counterclockwise, but qualitatively, there are two kinds of orbits, ultimately both converging to the stable fixed point. Subthreshold orbits make only small excursions; however, for trajectories starting with either low inactivation (roughly $W < -.45$) or high voltage (for example, $V > 0$), the system makes a large excursion around the right branch of the V nullcline, corresponding to a spike, before settling back to the fixed point. Although

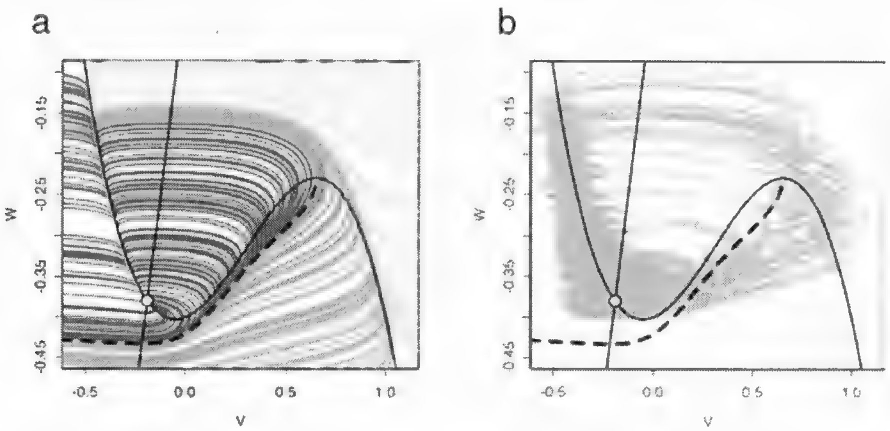


Figure 1: (a) Phase plane of the FitzHugh-Nagumo neuron model with trajectories and threshold (dashed line). Parameters are $a = 0.1$, $b = 0.5$, $c = -0.4$, and $\psi = 0.015$. (b) A single extended trajectory on the FN phase plane driven by white noise input, showing multiple spikes. The threshold (dashed line) still makes an approximate boundary of the spiking and nonspiking portions of the trajectory (gray).

we will generally be considering a time-varying input $I(t)$, we will make heavy use of the $I = 0$ phase plane; one can regard this as representing the instantaneous flow if the input current is switched off. For $I = 0$, one can define a separation of spiking initial conditions from subthreshold conditions, marked in light versus dark gray in Figure 1a. This separation of behavior can be thought of as a threshold. Typically the threshold for spiking is taken to be a threshold in voltage only, but this phase-plane picture demonstrates that one should really consider the threshold to depend on both V and the hidden inactivation variable(s) (Izhikevich, 2006).

One can approximate this "dynamical" threshold by considering a minimal spiking trajectory. There is no absolute distinction between spiking and subspiking trajectories in type II neurons (Gerstner & Kistler, 2002) such as Hodgkin-Huxley and FitzHugh-Nagumo, so this choice is necessarily somewhat arbitrary. Starting with a point on the phase plane at the end of the spike plateau, one can eliminate time from the dynamical equations, solving parametrically for the curve passing through this point. Here, we obtained one such curve by numerically solving the equations from the local maximum of the cubic nullcline, given by $\frac{d}{dV} V(1 - V)(a + V) = 0$.

One of the issues we explore in this article is the relationship between the geometry of the threshold in the multidimensional space of the dynamical variables and the measurements one makes in white noise analysis.

1.2 White Noise Analysis. Our approach is based on the functional perturbation expansion method, which we briefly review here. The simplest such expansion is known as the Volterra series (Volterra, 1930; Marmarelis,

& Marmarelis, 1974; Marmarelis, 2004), where the functional $y(t) = y[I(t)]$ is expanded as

$$\begin{aligned} y(t) = {} & h_0 + \int dt_1\, h_1(t_1) I(t-t_1) \\ & + \frac{1}{2!} \int dt_1 dt_2\, h_2(t_1, t_2) I(t-t_1) I(t-t_2) + \cdots, \end{aligned} \tag{1.7}$$

where the kernels h_i are called Volterra kernels. One problem with Volterra analysis is that the successive order kernels are not independent. For example, h_2 contains a constant component $\sim \int dt_1\, h_2(t_1, t_1)$.

Wiener analysis (Wiener, 1958; Rieke et al., 1997) avoids this issue. In this framework, $I(t)$ is assumed to be zero mean gaussian white noise satisfying

$$\langle I(t) I(t') \rangle = \sigma^2 \delta(t - t').$$

Under Wiener analysis, $y[I(t)]$ is expanded as

$$\begin{aligned} y(t) = {} & g_0 + \int dt_1\, g_1(t_1) I(t-t_1) \\ & + \frac{1}{2!} \left\{ \int dt_1 dt_2\, g_2(t_1, t_2) I(t-t_1) I(t-t_2) - \sigma^2 \int dt'\, g_2(t', t') \right\} + \cdots. \end{aligned} \tag{1.8}$$

In the Wiener expansion, each kernel is independent of the others, since at each order, one subtracts out the lower-order components. Therefore, the kernels g_i can be easily obtained by measuring the correlation functions between $y(t)$ and $I(t)$—for example,

$$\begin{aligned} \langle y(t) \rangle = {} & g_0 + \int dt_1\, g_1(t_1) \langle I(t-t_1) \rangle \\ & + \frac{1}{2!} \left\{ \int dt_1 dt_2\, g_2(t_1, t_2) \langle I(t-t_1) I(t-t_2) \rangle \right. \\ & \left. - \sigma^2 \int dt'\, g_2(t', t') \right\} + \cdots \\ = {} & g_0, \\ \langle y(t) I(t-t_1) \rangle = {} & g_1(t_1)/\sigma^2, \\ \langle y(t) I(t-t_1) I(t-t_2) \rangle = {} & g_2(t_1, t_2)/\sigma^4, \end{aligned}$$

and so on. The correlation method can be used to determine the relationship between Wiener and Volterra kernels. From equation 1.8,

$$\begin{aligned} g_0 &= h_0 + \int dt_1\, h_1(t) \langle I(t-t_1) \rangle \\ &\quad + \frac{1}{2!} \int dt_1 dt_2\, h_2(t_1, t_2) \langle I(t-t_1) I(t-t_2) \rangle + \cdots \\ &= h_0 + \frac{\sigma^2}{2!} \int dt'\, h_2(t', t') + \cdots, \\ g_1(t) &= h_1(t) + \frac{\sigma^2}{2!} \int dt' h_3(t', t', t) + \cdots, \end{aligned}$$

and so on.

Wiener analysis has been successfully used to determine the relevant feature space for neural systems by reverse correlation of the spiking output with white noise input. For this application, the output is taken to be the sequence of spike times, $\rho(t) = \sum_i \delta(t - t_i)$. Let us define the stimulus $\mathbf{s}$ as a linear transformation of the current input $I(t)$,

$$s_i = \int dt'\, f_i(t') I(t - t'), \tag{1.9}$$

where f_i are some set of independent linear filters. The goal is to determine how a single spike is generated by I, reduced to a vector $\mathbf{s}$ comprising projections along the relevant dimensions f_i.

We need to compute the probability of spiking as a function of the input, $P\,(\text{spike}|\mathbf{s})$,

$$P\,(\text{spike}|\mathbf{s}) = P(\text{spike}) G(\mathbf{s}), \tag{1.10}$$

where $G(\mathbf{s})$ is the input-output function relating $\rho(t)$ to the time-varying stimulus $\mathbf{s}(t)$, and we have separated out the scale factor $P(\text{spike})$ proportional to the mean firing rate. In principle, since the equations of motion are deterministic, $P\,(\text{spike}|\mathbf{s})$ is a Boolean function. However, as we will be approximating $\mathbf{s}$ by a finite-dimensional truncation and will be solving for P itself in successive orders, the result will no longer be Boolean.

We can compute $P\,(\text{spike}|\mathbf{s})$ using Bayes' theorem (Rieke et al., 1997),

$$G(\mathbf{s}) = \frac{P\,(\text{spike}|\mathbf{s})}{P(\text{spike})} = \frac{P\,(\mathbf{s}|\text{spike})}{P(\mathbf{s})}. \tag{1.11}$$

This shows that the Wiener kernels of $G(\mathbf{s})$ are the moments of $P\left(\mathbf{s}|\text{spike}\right)$, which can be easily measured. In this article, we focus on the first two moments of this distribution.[1] The first moment is the spike-triggered average (STA), $\bar{s}(t)$,

$$\bar{s}(t) = \left\langle I(t_i - t)\right\rangle_i = \int d\mathbf{s}\ \mathbf{s}P\left(\mathbf{s}|\text{spike}\right) = \int d\mathbf{s}\ \mathbf{s}P(\mathbf{s})G(\mathbf{s}), \tag{1.12}$$

where the average $\langle\cdot\rangle_i$ over the spike occurrence times $\{t_i\}$ is replaced with an average over the spike-triggering ensemble. When the neuron has multiple characteristic features, the STA may contain little information about them since it points in only a single direction, which is a certain linear combination of relevant features. The second-order kernel, however, can give such information. This kernel is related to the spike-triggered covariance (STC) matrix C, which we will define as

$$\begin{aligned} C(t,t') &= \left\langle I(t_i - t) - \bar{s}(t), I(t_i - t') - \bar{s}(t')\right\rangle_i - \left\langle I(\tau - t)I(\tau - t')\right\rangle_\tau \\ &= \sigma^2 g_0 \delta(t,t') + \sigma^4 g_2(t,t') - \bar{s}(t)\bar{s}(t') - \sigma^2\delta(t,t') \\ &= \sigma^4 g_2(t,t') - \bar{s}(t)\bar{s}(t'). \end{aligned} \tag{1.13}$$

$g_2(t,t')$ is the second-order Wiener kernel of $G(\mathbf{s})$. When a functional $F[\mathbf{s}]$ depends on gaussian noise $\mathbf{s}$, $\left\langle s_i F[\mathbf{s}]\right\rangle = \sum_j \left\langle s_i s_j\right\rangle \left\langle \partial_j F[\mathbf{s}]\right\rangle$. Thus, we can expand g_2 (Bialek & de Ruyter van Steveninck, 2005),

$$g_2(t,t') = \sum_{i,j} \left\langle \frac{\partial^2 G}{\partial s_i \partial s_j}\right\rangle f_i(t) f_j(t'), \tag{1.14}$$

where equation 1.9 introduces the linear filters that define the components of $\mathbf{s}$.

Therefore, unless the modified Hessian $\hat{H}_{ij} = \langle \frac{\partial^2}{\partial s_i \partial s_j} G\rangle - \langle \frac{\partial}{\partial s_i} G\rangle\langle \frac{\partial}{\partial s_j} G\rangle$ has null directions,[2] the STC is spanned by the relevant feature space $\{f_i\}$. Diagonalizing the STC and extracting the eigenvectors with the largest absolute eigenvalue will determine the leading dimensions spanning the feature space (Brenner et al., 2000; Agüera y Arcas et al., 2003; Bialek & de Ruyter van Steveninck, 2005). Eigenmodes appearing with negative eigenvalue correspond to stimulus directions in which the variance of the spike-conditional distribution is reduced relative to the stimulus prior distribution; eigenmodes with positive eigenvalues are directions with increased variance. Such cases may arise if the spike-conditional distribution is bimodal in some direction or if it forms a ring around the origin; phase-invariant (complex) cells are an example (Simoncelli, Pillow, Paninski, & Schwartz, 2004; Rust et al., 2005; Fairhall et al., 2006; Petersen, 2003; Maravall et al., 2007).

---

[1] Note that the zeroth-order kernel is simply $g_0 = 1$.

[2] Since $\bar{s}(t)\bar{s}(t')$ is only rank one, there are only two possibilities leading to null directions. The first is when $\langle \partial_i \partial_j G\rangle$ has null directions, in which case it is excluded (for our purposes) by definition. The second case is when $\langle \partial_i \partial_j G\rangle$ has an eigenvector $\bar{s}(t)$ with an eigenvalue $||\bar{s}\,(t)||^2$, which does not happen generically.

In this article, we work with a multidimensional space of dynamical variables, as in equations 1.1 to 1.6, for which we obtain the Volterra series. When we denote them with y_i, the Wiener kernels for $G(y_i)$ can be expanded as

$$\begin{aligned} g_1[G] &= \sum_i \frac{\partial G}{\partial y_i} g_1[y_i] + \cdots, \\ g_2[G] &= \sum_i \frac{\partial G}{\partial y_i} g_2[y_i] + \sum_{i,j} \frac{\partial^2 G}{\partial y_i \partial y_j} g_1[y_i] g_1[y_j] + \cdots \\ &\cdots \end{aligned}$$

where $g_n[y_i]$ is an nth Wiener kernel of y_i and dots represent higher-order terms than quadratic. Here we note that $g_n[y_i]$ is generated by the infinite series of Volterra kernels. Moreover, we will see later that the Volterra kernels at each order can be constructed from the set of first-order kernels.

2 Linear Systems

In section 1, we discussed the application of white noise methods to experimental data where the output of the system is measured as the firing times of spikes. With access to the entire dynamical system, one could attempt to model the relationship between the input $I(t)$ and the output voltage $V(t)$, including both subthreshold and spiking behavior. As we have discussed, neurons are highly nonlinear. Perturbative expansions such as Wiener and Volterra series generally are unable to capture the nonlinearities of spike generation with few terms. However, a very successful and highly influential approach to the modeling of neural systems has been to separate a linear filtering stage from an explicit nonlinearity, capturing the very sharp voltage increase of the spike (Victor & Shapley, 1980; Gerstner & Kistler, 2002; Kistler, Gerstner, & van Hemmen, 1997; Berry & Meister, 1999). This is known as a cascade model, and in various forms is the theoretical basis for many of the simple neuron models in use, including integrate-and-fire neuron (Gerstner & Kistler, 2002; Agüera y Arcas & Fairhall, 2003), the spike response model (Gerstner & Kistler, 2002), generalized integrate-and-fire models (Paninski, Pillow, & Simoncelli, 2004) and cascade models of retinal ganglion cells (Victor & Shapley, 1979, 1980; Keat, Reinagel, Reid, & Meister, 2001; Pillow, Paninski, Uzzell, Simoncelli, & Chichilnisky, 2005).

The success of these models suggests that a breakdown of the dynamical system into its linear component and a threshold is a useful approximation to examine. In our analysis of the dynamical system, we will take the approach of experimental work, where the output is reduced to a sequence of spike times. Our goal here is to derive the form of the multidimensional linear stage that arises directly from the equations of motion and examine the consequences of thresholds of various forms.

We begin by walking through some simple linearized systems analysis and clarify the connection with the cascade model picture. Rewriting equations 1.1 to 1.4 in the following simple form,

$$\partial_t y_k = f_k(y_1, y_2, \ldots, y_N) + I(t)\delta_{k0}, \quad k = 0, 1, \cdots, n-1, \tag{2.1}$$

we assume that they have a fixed-point solution $y_k(t) = y_k^{(0)}$ when $I(t) = 0$. Now the linear approximation of the system around the fixed point is given by[3]

$$\mathcal{D}_{km} y_m^{(1)} = I(t)\delta_{k0}, \tag{2.2}$$

where

$$\mathcal{D}_{km} = \delta_{km}\partial_t - J_{km}, \quad J_{km} = \left.\frac{\partial f_k}{\partial y_m}\right|_{y^{(0)}}. \tag{2.3}$$

The linear filters are the kernels of equation 2.3, $\epsilon^{(1)} = \mathcal{D}^{-1}$. They satisfy

$$\mathcal{D}_{km}\epsilon_{ml}^{(1)}(t) = \delta_{kl}\delta(t), \tag{2.4}$$

giving the linear filter form for the system's evolution:

$$y_k^{(1)}(t) = \int_{-\infty}^{t} dt'\, \epsilon_k^{(1)}(t-t')I(t'), \tag{2.5}$$

where $\epsilon_k^{(1)} = \epsilon_{k0}^{(1)}$. Equation 2.4 can be solved by diagonalizing the Jacobian matrix J_{km}. Letus consider the following diagonalization with eigenvalues λ_α:

$$\lambda_\alpha \delta_{\alpha\beta} = U_{\alpha k} J_{km} U_{m\beta}^{-1}. \tag{2.6}$$

---

[3] From here on, we use the Einstein summation convention, for example, $C_{km}d_m \equiv \sum_m C_{km}d_m$.

Here we employ the notation that indices in the diagonalized basis are Greek letters. Therefore, equation 2.4 can be written in the new basis as

$$\mathcal{D}_\alpha \tilde{\epsilon}_\alpha^{(1)} = U_{\alpha 0}\delta(t), \quad \mathcal{D}_\alpha = \partial_t - \lambda_\alpha, \tag{2.7}$$

where the first-order kernels $\tilde{\epsilon}^{(1)}$ in this basis are

$$\tilde{\epsilon}_\alpha^{(1)}(t) = e^{\lambda_\alpha t} H(t),$$

where $H(t)$ is the Heaviside step function and arises from causality. The eigenvalues λ_α can be real or complex. The sign of the real part of the eigenvalues is determined by the stability properties of the fixed point at $y^{(0)}$. For a neural system, this fixed point should be attracting, so all eigenvalues have negative real part, and thus eigenvectors decay as $t \to \infty$. This underscores that the linear filters can take only a limited number of forms: decaying exponentials or damped oscillations.

We note the following facts. First, in general, an n-dimensional system will have n independent first-order kernels, except when the Jacobian has a degenerate spectrum. Further, the linear kernels can be generated from a "master kernel,"

$$\varphi(t) = \sum_\alpha e^{\lambda_\alpha t} H(t),$$

and its time derivatives up to nth order. To show this, from the expression of the derivative of a master kernel,

$$\partial_t{}^k \varphi(t) = \sum_\alpha \lambda_\alpha^k e^{\lambda_\alpha t} = T_{k\alpha} e^{\lambda_\alpha t}, \quad T_{k\alpha} = \lambda_\alpha^k, \quad (k, \alpha = 0, \ldots, n-1),$$

up to additive singular terms coming from the derivatives of $H(t)$. $T_{k\alpha}$ is a Wronskian matrix of exponentials at $t = 0$. The determinant of $T_{k\alpha}$ is the Vandermonde determinant $\det(T_{k\alpha}) = \prod_{\alpha > \beta}(\lambda_\alpha - \lambda_\beta)$, which cannot vanish by definition (Gradshteyn & Ryzhik, 2000). This means that the master kernel and its derivatives can be transformed to $e^{\lambda_\alpha t}$'s in a nonsingular way.

Therefore, we conclude that all first-order kernels can be expressed in terms of a master kernel and its time derivatives. Frequently the first-order kernel is nonzero at $t = 0$, so that the time derivative of the filter also includes a delta function at $t = 0$. We will see the appearance of this singular component in the white noise analysis.

To recover the filters corresponding to the dynamical variables, one must invert the diagonalizing operation, equation 2.6, to obtain linear combinations of the eigenmodes:

$$\epsilon_i^{(1)}(t) = U_{i\alpha}^{-1}\tilde{\epsilon}_\alpha^{(1)}(t) = \sum_\alpha c_\alpha e^{\lambda_\alpha t} H(t), \tag{2.8}$$

where the coefficients c_α derive from the components of the Jacobian. Thus, the filters of the linear system are sums of exponentials—either purely real or with an imaginary component. To the extent that subthreshold dynamics are well approximated by the linearized system, this shows clearly why one might expect to find both integrate-and-fire-like neurons, with filters having purely real associated eigenvalues, and oscillate-and-fire neurons (Izhekevich, 2001; Hutcheon & Yarom, 2000) with a corresponding eigenvalue having a nonzero imaginary part. In terms of the linear system, the possible set of filters or feature space of the system is dual with the dynamical variables that define the system. The subset of features that are relevant to a spike occurring are those that contribute to crossing the internal state over "threshold."

The properties outlined above have an interesting implication for white noise analysis in the case that the system is well described by subthreshold linearity. As we will see in the next section, the STA is typically a linear combination over the relevant features and their derivatives. If the components of the STA belong to the true feature space, as they typically will, one can derive the basis for the feature space from the STA and its successive order time derivatives, as has been empirically observed (Rieke, 1991).

2.1 Higher-Order Series Terms. The higher-order approximations to the system can also be expressed in terms of first-order kernels. For example, the second-order approximation is given by the following equation:

$$\mathcal{D}_{km} y_m^{(2)} = \frac{1}{2!} H_{kmn} y_m^{(1)} y_n^{(1)}, \qquad H_{kmn} = \left. \frac{\partial^2 f_k}{\partial y_m \partial y_n} \right|_{y^{(0)}}, \tag{2.9}$$

whose solution is

$$\begin{aligned} y_k^{(2)} &= \frac{1}{2!} \int ds_1 ds_2\, \epsilon_k^{(2)}(t - s_1, t - s_2) I(s_1) I(s_2), \\ \epsilon_k^{(2)}(t_1, t_2) &= \int dt'\, H_{lmn} \epsilon_{kl}^{(1)}(t') \epsilon_m^{(1)}(t_1 - t') \epsilon_n^{(1)}(t_2 - t'). \end{aligned} \tag{2.10}$$

In this way, all of the higher-order kernels can be obtained by outer products of first-order kernels. The detailed computation is summarized in appendix A.

2.2 White Noise Analysis of Threshold-Crossing Linear Neurons. In this section, we assume that the higher-order nonlinearity of the system is well captured by a threshold. Since voltage is generally the only observed variable, "the threshold" is often taken to be a threshold on the voltage. The imposition of a threshold on the voltage alone has been applied even to neural models such as FN and the Hodgkin-Huxley model system, where there is no difficulty in defining a threshold in V and the "hidden" dynamical variables of W (FN) or m, n, and h (HH). Here we consider the implications of considering the threshold to apply in multiple dimensions. Before we discuss the validity of the threshold approximation for an FN neuron, we illustrate the results from applying white noise analysis to models with a variety of threshold structures acting on subthreshold linear dynamics given by the linearized FN system.

Previous work has treated simple one-dimensional linear and nonlinear models, in which a spike is defined when the output of a single linear filter f on a random current input crosses a threshold. One can show (DeWeese, 1995; M. Meister, private communication, November 2004) that covariance analysis on this model finds two modes: the filter f and its time derivative, f'. The derivative mode is a result of the criterion defining the spike only on the upward crossing of the threshold:

$$\frac{d}{dt}\int^{t} d\tau f(t-\tau)I(\tau) > 0 \quad \rightarrow \quad \int^{t} d\tau f'(t-\tau)I(\tau) > 0. \tag{2.11}$$

Thus, the variance of projections onto f' is also reduced with respect to the prior. Covariance analysis has also been performed (Fairhall et al., 2006) on more realistic models in which an exponentially decaying afterhyperpolarization (AHP) is added to the voltage following a spike (Gerstner & Kistler, 2002; Keat et al., 2001). This makes the threshold-crossing model more realistic, but the manipulation does not alter the (f, f') eigenmode structure unless the timescale of the AHP is short enough that spiking occurs many times during the superthreshold fluctuation, which destroys the correlation with positive f'.

Here we assume that the system evolves according to linear subthreshold dynamics. We treat five cases. In the first three, the threshold is linear, but in different components of the 2D space; in the last two, the threshold is a nontrivial two-dimensional structure. The thresholds used are:

1. A traditional threshold in V only
2. A threshold in W only
3. A threshold on a linear combination of V and W
4. The union of two piecewise linear segments in V and W, that is, a logical "or" on 1 and 2
5. A curved threshold on a smooth, nonlinear function of V and W

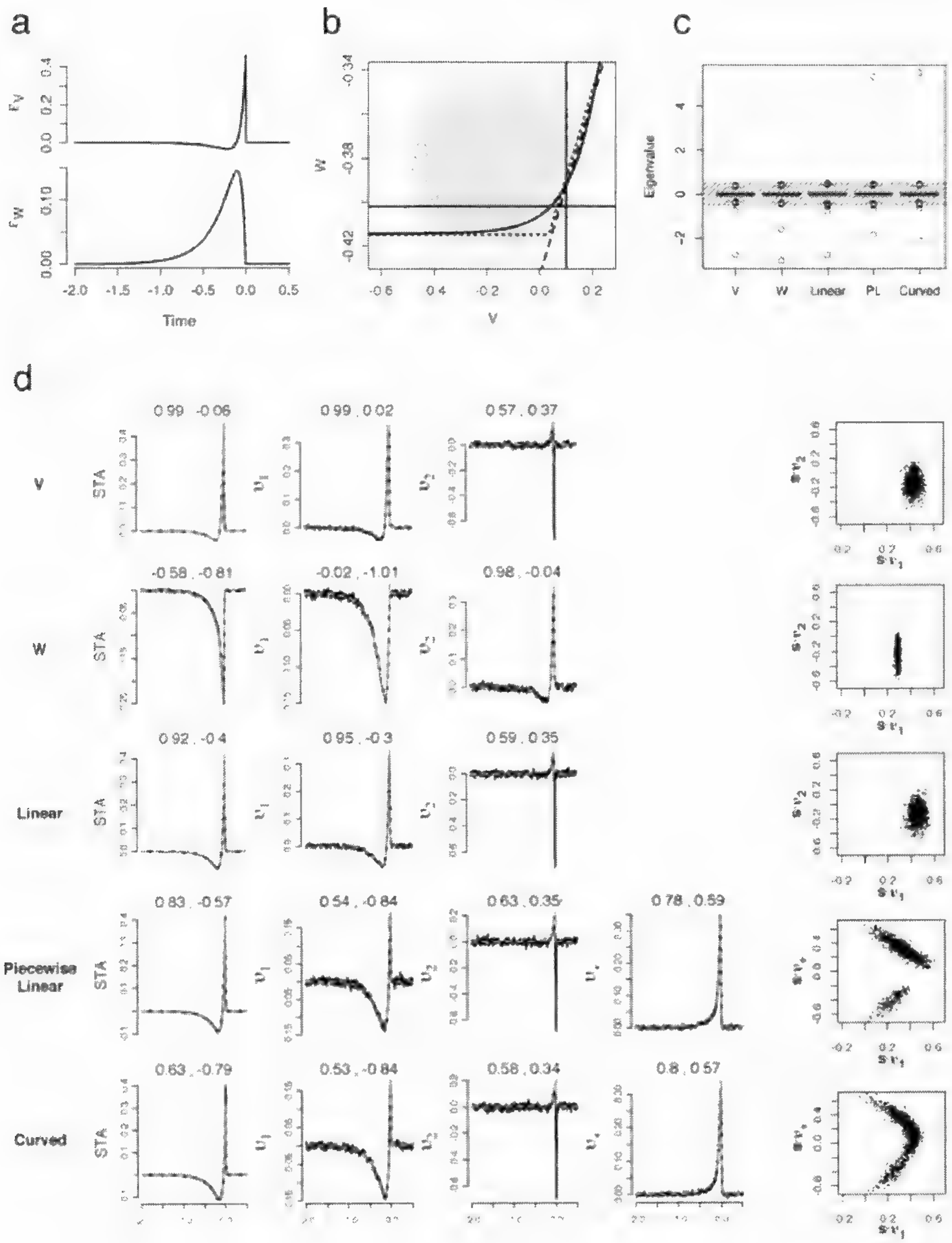
From the discussion in Section 2, the linear space in which the n-dimensional system operates is defined by the n filter inversions of the dynamical system, (see equation 2.5). The effect of the threshold is to select as relevant from those n dimensions the subspace defined by the filter directions that have a nonzero projection onto the threshold. Thus, we would predict that for the linear threshold, we will find two relevant modes: the primary direction f, which is normal to the direction of the threshold, and f', its time derivative. For a threshold in a single dimension, V or W only, f should recover the V or W filter, respectively, and the corresponding time derivative. For a threshold that is a linear combination of V and W, f should be a linear combination of the V and W filters. The distribution of trajectories is simply a linear transformation of the original gaussian stimulus distribution. Any linear threshold in this plane will produce negative eigenmodes, as the set of points selected by the threshold must have decreased variance with respect to this prior.

When the threshold is not linear in V and W, we expect to find two primary filters, f_1 and f_2, which will be some linear combination of the V and W filters, and, in principle, the time derivatives. However, the number of modes will be less than or equal to $n+1$ because, as we have shown, the time derivatives span the same space as the linear variables themselves. (Note that the $+1$ accounts for the singular component of the time derivative, $\delta(t)$, arising from discontinuity of the filters at $t = 0$.) Furthermore, a threshold that is not linear may produce a spike-conditional distribution with a direction in which the variance is actually increased with respect to the prior; such a direction will appear with a positive eigenvalue.

Figure 2b shows the thresholds that we applied to the FN first-order linear dynamics. The results of the covariance analysis are shown in Figures 2c and 2d. Figure 2c shows the covariance eigenvalue spectra for all five cases. The significant eigenvalues are empty circles beyond the error level denoted by the shaded box.[4] In the cases with thresholds on V only, on W only, and linear in both V and W, only two modes are obtained. With 2D thresholds, cases 4 and 5, three eigenmodes appear. Figure 2d shows the corresponding modes. We see that the V threshold picks out $\epsilon_V^{(1)}$, the W threshold selects $\epsilon_W^{(1)}$, and the (V, W)-linear threshold leads to a linear

---

[4] Significance of eigenvalues can be evaluated as follows. Let N be the number of spikes and d the dimension of each sampled stimulus. To estimate the finite size or dimension error, we choose N d-dimensional stimuli at random (Rust et al. 2005). In this case, the covariance matrix C in equation 1.14 is filled with $d(d+1)/2$ random numbers drawn from a normal distribution $\mathcal{N}(0, \sigma^2\sqrt{2/N})$ in the large N limit. When $d \gg 1$, the eigenvalues of this matrix follow Wigner's semicircle law, and their distribution is bounded by $\pm 2\sqrt{2}\sigma^2\sqrt{d/N}$, which is the error level (Mehta, 1991). When the stimulus has small correlation, the semicircle is not a good approximation of the eigenvalue distribution, but change in the upper and lower bounds rarely exceeds an order of magnitude, which can be checked by numerical simulation.


a
b
c
r_V
r_W
Time
-2.0
-1.5
-1.0
-0.5
0.0
0.5
W
V
-0.34
-0.38
-0.42
-0.6
-0.4
-0.2
0.0
0.2
Eigenvalue
V
W
Linear
PL
Curved
d
V
W
Linear
Piecewise Linear
Curved
STA
v_1
v_2
v_4
0.99, -0.06
0.99, 0.02
0.57, 0.37
-0.58, -0.81
-0.02, -1.01
0.98, -0.04
0.92, -0.4
0.95, -0.3
0.59, 0.35
0.83, -0.57
0.54, -0.84
0.63, 0.35
0.78, 0.59
0.63, -0.79
0.53, -0.84
0.58, 0.34
0.8, 0.57
$s \cdot v_1$
$s \cdot v_2$
$s \cdot v_4$

combination of these filters. In the latter two cases, the eigenmodes are a linear combination of $\epsilon_V^{(1)}$, $\epsilon_W^{(1)}$, and $\partial_t \epsilon_V^{(1)}$, itself a linear combination of the V and W kernels and $\delta(t)$.

Figure 2d also shows the projection of spike-triggered stimuli onto two of the distinguished eigenvectors. In the first three cases, the linearity of the threshold is manifested in the 2D plane as a single elliptic cluster, while the nonlinear threshold cases show richer structure.

Recall that the filters are generated by linear combinations of a master filter and its time derivatives. This implies that our multidimensional threshold structure has a natural interpretation as a dynamical threshold, depending not only on the voltage V but also on $\partial_t V$, $\partial_t{}^2 V$, and so on (Azouz & Gray, 2000). We will show how well this approximation fits for the Hodgkin-Huxley model in the final section.

2.3 Analysis of Threshold-Crossing Models in Multiple Dimensions.

An advantage of the simple models introduced in the previous section is that they can be treated analytically. From equation 1.11, we have

$$G(\mathbf{s}) = \frac{P\left(\text{spike}|\mathbf{s}\right) P(\mathbf{s})}{P(\text{spike})} = \frac{P\left(\text{spike}|\mathbf{s}\right) P(\mathbf{s})}{\int D\mathbf{s}\, P\left(\text{spike}|\mathbf{s}\right) P(\mathbf{s})},$$

where $\int D\mathbf{s} = \int ds_1 ds_2 \cdots$. Instead of directly computing $G(\mathbf{s})$, we characterize it by a moment generating function:

$$W[\mathbf{j}] = \log \int D\mathbf{s} P(\text{spike}|\mathbf{s}) P(\mathbf{s}) e^{\mathbf{j}\cdot\mathbf{s}}. \tag{2.12}$$

We define an n-dimensional first-order system, given by dynamical variables y_k and the corresponding linear filters ϵ_k as in equation 2.2. We assume

---

Figure 2: (a) First-order kernels of the FN neuron model, drawn as filters, that is, in $-t$. (b) Various thresholds for the linear system. V (vertical), W (horizontal), linear (dashed), piecewise linear (dotted), and curved (solid curve) thresholds are used. (c) Spectra of covariance matrices for the thresholds. The shaded box represents the level of error from finite sample size and dimension. (d) STAs and covariance modes of the linear system with various thresholds. The nonsingular part of each mode is fitted using least squares to a linear combination of the (orthogonalized) first-order kernels as $v_{fit} = c_V \epsilon_V^{(1)} + c_W \epsilon_W^{(1)}$. The kernels are normalized, excluding the δ-function component. The coefficients (c_V, c_W) are displayed above each plot, and the gray line is a fitted function. For each case, we show the projection of spike-triggered current histories onto either the leading two negative modes, v_1 and v_2, or v_1 and v_+, depending on the threshold.

that the stimulus is a gaussian white noise current $I(t)$ with zero mean and variance σ^2. Thus, as before,

$$y_k(t) = \int_0^\infty d\tau \epsilon_k(\tau) I(t-\tau).$$

We denote a random segment of the current stimulus $I(\tau)$, $\tau \le t$ as an infinite-dimensional vector $\mathbf{s}$.[5] Any sample of $\mathbf{y}$ is therefore a functional of the random variable $\mathbf{s}$, $\mathbf{y}[\mathbf{s}]$; for simplicity, we will write $\mathbf{y}$.

Spiking is determined by crossing a threshold $\theta(\mathbf{y}) = 0$ in the phase space from below. Then,

$$P\left(\text{spike}|\mathbf{s}\right) = \delta(\theta(\mathbf{y})) H(\dot{\mathbf{y}} \cdot \nabla_{\mathbf{y}} \theta(\mathbf{y})) \dot{\mathbf{y}} \cdot \nabla_{\mathbf{y}} \theta(\mathbf{y}).$$

The Heaviside function $H(\cdot)$ ensures that spiking occurs only on a threshold crossing from below, and the weight factor $\dot{\mathbf{y}} \cdot \nabla_{\mathbf{y}} \theta(\mathbf{y})$ is a geometric factor accounting for the flux at the threshold. The filters ϵ_k are not necessarily normalized or orthogonal to each other, and it is convenient to define the orthonormal basis $\{f_\mu\}$, which spans the same space as $\{\epsilon_k\}$. Then there will be a linear transformation $T_{\mu k}$,

$$f_\mu = T_{\mu k} \epsilon_k,$$

and so we define a new coordinate system:

$$z_\mu \equiv T_{\mu k} y_k.$$

As $\mathbf{s}$ is gaussian, the orthonormally transformed variable $\mathbf{z}$ is also gaussian with variance σ^2.

Now we can separate the stimulus into two components: its projection into the subspace spanned by the $\{f_\mu\}$ and the orthogonal component. We denote the corresponding directions of $\mathbf{j}$ as $\mathbf{j} = \mathbf{j}_\parallel + \mathbf{j}_\perp$, where

$$j_{\parallel\mu}(t) = \int_0^\infty d\tau\, j(t-\tau) f_\mu(\tau).$$

---

[5] In practice, any application to data requires discretization in time. We use the convention that a function $f(t)$ is discretized as $\hat{f}_t = f(t)\sqrt{\Delta t}$, where Δt is the time step. This implies $\int f^2 dt \approx \sum_i f(i\Delta t)^2 \Delta t = \sum_i \hat{f}_i^2 = \hat{\mathbf{f}} \cdot \hat{\mathbf{f}}$, and thus the vector $\hat{\mathbf{f}}$ obtained by discretizing $f(t)$ has a vector norm corresponding to the L^2 norm.

The moment generating function can then be separated as $W[\mathbf{j}] = W_\parallel[\mathbf{j}_\parallel] + W_\perp[\mathbf{j}_\perp]$, where

$$W_\perp[\mathbf{j}_\perp] = \log \int D\mathbf{s}_\perp\, e^{-s_\perp^2/2\sigma^2 + \mathbf{j}_\perp\cdot\mathbf{s}_\perp} = \frac{\sigma^2}{2} j_\perp(t)^2 + \text{const}, \tag{2.13}$$

where $\mathbf{s}_\perp$ are the components of $\mathbf{s}$ orthogonal to the plane spanned by $\mathbf{z}$, and

$$\begin{aligned} W_\parallel[\mathbf{j}_\parallel] &= \log \int d^n\mathbf{z}\, \det(T^{-1}) e^{-z^2/2\sigma^2}\delta(\theta(\mathbf{z}))H(\dot{\mathbf{z}}\cdot\nabla_{\mathbf{z}}\theta(\mathbf{z}))\dot{\mathbf{z}}\cdot\nabla_{\mathbf{z}}\theta(\mathbf{z})e^{\mathbf{j}_\parallel\cdot\mathbf{z}} \\ &= \log \int d^n\mathbf{z}\, e^{-z^2/2\sigma^2}\delta(\theta(\mathbf{z}))w(\mathbf{z})e^{\mathbf{j}_\parallel\cdot\mathbf{z}} + \text{const}, \end{aligned} \tag{2.14}$$

with

$$w(\mathbf{z}) = H((R\mathbf{z})\cdot\nabla_{\mathbf{z}}\theta(\mathbf{z}))(R\mathbf{z})\cdot\nabla_{\mathbf{z}}\theta(\mathbf{z}),$$

where $R = TJT^{-1}$ is the Jacobian matrix of the linear system, equation 2.2, in the $\{f_\mu\}$ basis, and we use

$$\dot{z}_\mu = T_{\mu k}\dot{y}_k = T_{\mu k}J_{kl}y_l = (TJT^{-1})_{\mu\nu}z_\nu.$$

As previously discussed, the closure under time differentiation of the space spanned by the first-order kernels is assured by $\dot{\mathbf{z}}$, as in equation (A.5) in appendix A. Note that our separation of $W[\mathbf{j}]$ depends on this particular property. For a fixed spike time t, restricting ourselves to the region where $\tau < t$, $W_\perp[\mathbf{j}]$ becomes a static integral. Thus, equation (2.14) shows clearly how the properties of this model emerge. P (spike|$\mathbf{s}$) is given by a threshold with a weight function $w(\mathbf{z})$ up to some linear transformations. Therefore, for example, the linear threshold case reduces to the one-dimensional case. Since $\nabla_{\mathbf{z}}\theta(\mathbf{z})$ is a constant vector, every computation reduces to a one-dimensional integral in this direction, and the corresponding filter is a linear combination of f_μs, as we have observed in Figure 2d. Furthermore, a variety of structures may be generated depending on how trajectories cross the threshold.

From equations 2.14 and 2.13, we can derive the analytic forms of the first- and second-order moments:

$$\begin{aligned} &\mathrm{STA}(\tau) = \sum_\mu c_\mu f_\mu(\tau), \qquad c_\mu = \frac{1}{N}\int d^n\mathbf{z}\, z_\mu e^{-z^2/2\sigma^2} w(\mathbf{z}), \\ &N = \int d^n\mathbf{z}\, e^{-z^2/2\sigma^2} w(\mathbf{z}), \end{aligned} \tag{2.15}$$

and

$$C(\tau,\tau') = \sum_{\mu,\nu} \left(c_{\mu\nu} - c_\mu c_\nu - \sigma^2 \delta_{\mu\nu}\right) f_\mu(\tau) f_\nu(\tau'),$$
$$c_{\mu\nu} = \frac{1}{N} \int d^n\mathbf{z}\, z_\mu z_\nu e^{-z^2/2\sigma^2} w(\mathbf{z}). \tag{2.16}$$

Note that we have assumed a prior based on the distribution of randomly driven trajectories of the linear system. This takes no account of perturbations of this distribution due to flux from the system's return from a spike. This assumption is equivalent to assuming that the last spike is in the distant past, so that memory of that perturbation has vanished. In this article, we consider only this "isolated spike" case. We return to this point in the discussion.

2.4 Variance Dependence. Through equations 2.15 and 2.16, this model captures an explicit dependence of the white noise outcome on the stimulus statistics, in particular, the variance σ^2. While the subthreshold dynamics are linear, the dependence on σ is nonlinear due to the threshold shape and the weight function $w(z)$. We discuss two examples below.

We first consider a linear threshold. Through a suitable linear transformation, this case simply reduces to a filter-and-fire model with a single filter, say, f_0, and a fixed threshold $z_0 = \theta_f$. Now the distribution of threshold crossing points is constrained by $z_0 = \theta_f$ and $\dot{z}_0 > 0$. As we mentioned in the previous section, $\dot{z}_0$ lies in the originally defined feature space and is given by another single filter, which we denote by f_1. In other words, when the normalized $\dot{f}_0$ is denoted by f_0', we can choose $f_1 = f_0'$ by a suitable orthogonal linear transformation. Thus, our system depends on two filters, which are depicted schematically in Figure 3a.

The STA is the centroid of the distribution of threshold crossing points. Equation 2.15 reproduces a previously known result (M. Meister, private communication November 2004; Fairhall et al., 2006),

$$\mathrm{STA}(\tau) = \theta_f f_0(\tau) + \frac{\sigma}{\sqrt{\pi/2}} f_1(\tau). \tag{2.17}$$

With high-variance $\sigma \gg \theta_f$, the STA is dominated by f_1. In the feature space, with increasing variance, the threshold stays the same, but a larger portion of it is crossed by trajectories driven by the larger variance ensemble, as can be seen in Figure 3a.

When the threshold is curved, the σ dependence is considerably more complicated. We will consider an extreme but analytically tractable version

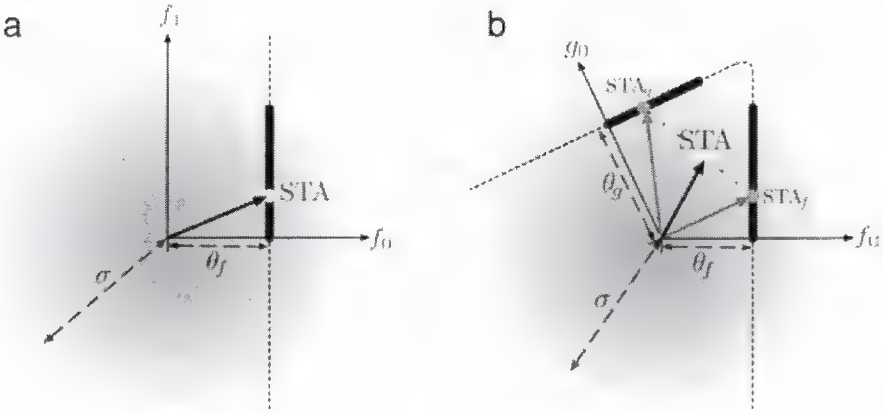


Figure 3: (a) Diagram of a single filter model. The disc of radius σ represents the prior gaussian distribution. The dotted line is the model threshold imposed at $z_0 = \theta_f$. The thick line is a distribution of spike-triggered stimuli, and the white dot represents its average, the STA. (b) Diagram of a model whose threshold is the union of two lines, each imposed at a distance $\theta_{f,g}$ from the center. Two gray dots denote the STAs for each segment.

of this case to illustrate the point: let the curved threshold be approximated by two linear ones imposed at θ_f and θ_g in the f_0 and g_0 directions, respectively, as in Figure 3b. In this case, one segment of the threshold imposes a dependence on the filters f_0 and its (normalized) derivative, f_1, while the other selects g_0 and $g_1 = \dot{g}_0/\|\dot{g}_0\|$. The space of relevant features is still two-dimensional or three-dimensional, including the δ-function, since g_0 and g_1 are linear combinations of f_0, f_1, and possibly also $\delta(t)$.

Now the STA of this system is

$$\begin{aligned} \mathrm{STA}(\tau) &= \cos^2\varphi \cdot \mathrm{STA}_f(\tau) + \sin^2\varphi \cdot \mathrm{STA}_g(\tau), \\ \varphi &= \tan^{-1} e^{(\theta_f^2-\theta_g^2)/4\sigma^2}, \end{aligned} \tag{2.18}$$

where $\mathrm{STA}_\epsilon(\tau) = \theta_\epsilon \epsilon_0(\tau) + \frac{\sigma}{\sqrt{\pi/2}}\epsilon_1(\tau)$ as in equation 2.17. Note that as in Figure 3b, the STA does not lie on the threshold. As for a variety of experimental examples such as complex cells (Touryan et al., 2002), neurons of rat barrel cortex (Petersen, 2003), and some retinal ganglion cells (Fairhall et al., 2006), the spike-triggered stimuli are poorly represented by their first-order statistics, the STA. Also, the coefficients of $\mathrm{STA}_{f,g}$ depend exponentially on $\theta_{f,g}$ and σ. Thus, the system shows a nonlinear dependence on the stimulus variance.

Figure 4 shows an example. The W threshold model does not show any significant change, except sharpening of the STA due to the increased component of $\dot{\epsilon}_W$, Figure 4b, and linear broadening of the spike-triggered stimulus distribution. The piecewise linear threshold case is more dramatic: while

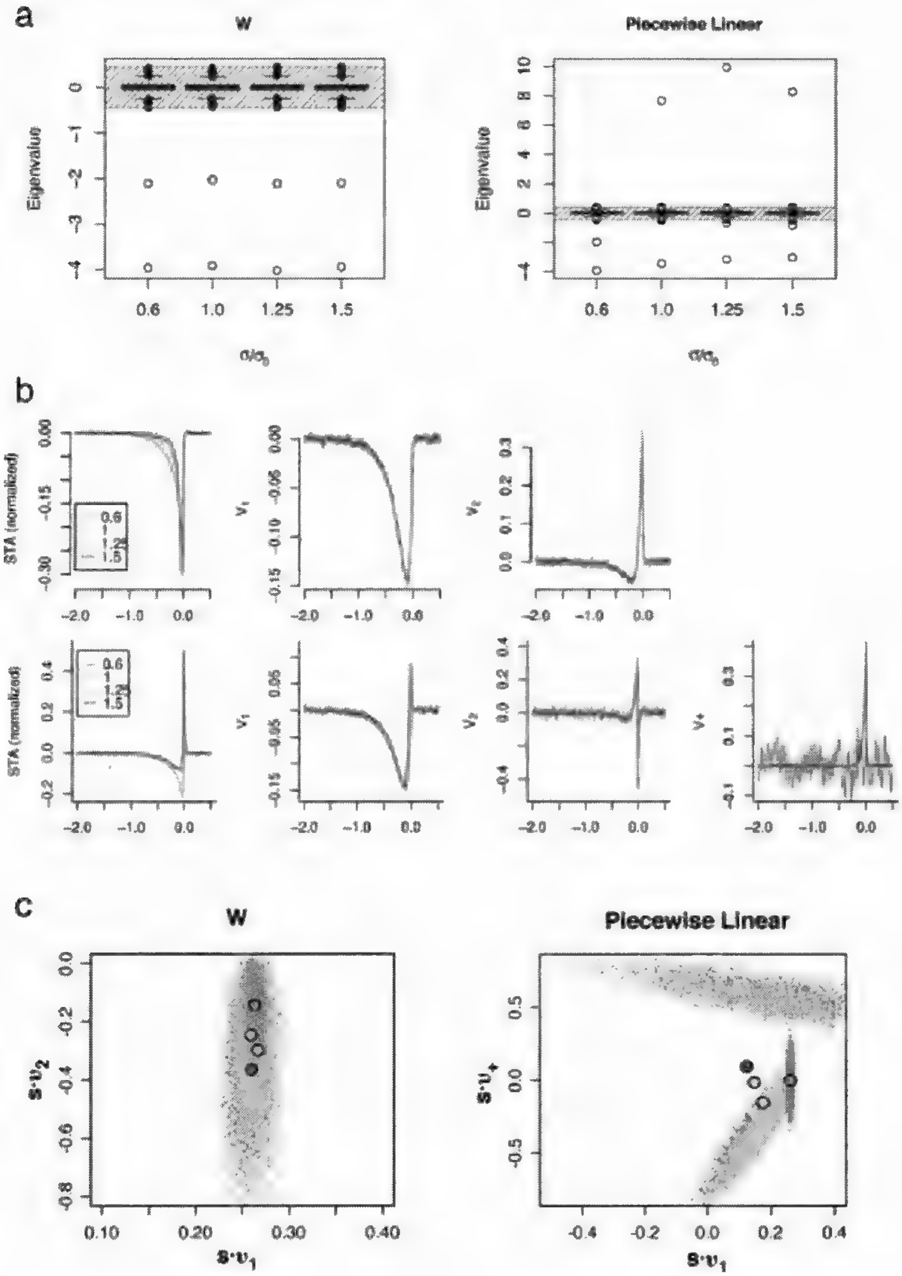


Figure 4: Two examples of variance dependence in the linear system model. (a) Spectra of the covariance matrix for several different variance stimuli. (b) Significant eigenmodes. (c) Projection of spike-triggering stimuli onto the filter subspace defined by $v_{1,2}$ for the linear and $v_{1,+}$ for the piecewise linear threshold, respectively, colored according to the stimulus variance as in the legend of b. The circled points are the centers of the distributions, corresponding to the respective STAs, colored according to the respective distribution.

the smallest variance does not drive the system hard enough to produce a positive eigenmode (hence, the one-dimensional distribution of projections seen in Figures 4b and 4c, red), a new significant mode emerges as the variance increases. The STA changes beyond sharpening and almost looks like a different model at high variance compared with low variance. Each of the significant modes changes as more trajectories cross the threshold from the other side and the principal axis of the distribution rotates, as seen in Figure 4c.

3 Dynamical Threshold

In the previous section, we considered neuron models composed of linear filters derived from the FN neuron, and some choices of imposed thresholds. However, when we consider the full FN neuron, the threshold arises from the structure of the FN equations, 1.5 and 1.6. Here we discuss the importance of the threshold identification for reverse correlation analysis.

The problem of the identification of a threshold arises immediately on attempting a reverse correlation analysis; spike times are often defined by the threshold crossing of the voltage. Here also we can impose an arbitrary threshold in V to identify each spike. The STA obtained using this scheme is displayed in Figure 5a. It is clear that it cannot be well described by the first-order kernels. However, this does not mean breakdown of the analysis; rather, it underscores the point that the STA or any other single spike quantities should be computed using the "correct" threshold that we take to be the dynamical threshold discussed in section 1.

Figure 5b compares the spike-triggered stimuli in the fixed V and dynamical threshold cases. While the peaks of the stimuli are spread out in time in the fixed V threshold case, for the dynamical threshold, the peaks lie in a narrow band around the spike time. As we can see from Figure 1, the spiking trajectories cross the dynamical threshold before crossing the V threshold, and this timing difference, Δt, depends on W. Additionally, since the system is driven by gaussian white noise, the variability increases with the timing difference, inducing a point-spread function on the STA. Hence, each spike-triggered stimulus is contaminated by this temporal jitter, and estimated filters are distorted. As mentioned in section 1.1, the FN neuron does not have a clear-cut dynamical threshold; the threshold was chosen to some degree arbitrarily, and this will induce some error in the estimation of filters. However, Figure 6a shows the improvement in the STA computed using the dynamical threshold. Figures 5c and 5d show the point-spread distribution—the distribution of values of Δt—and how it is correlated with W. This situation is similar to that discussed in Aldworth, Miller, Gedeon, Cummins, & Dimitrov, (2005), where it was noted that temporal or spatial jitter can blur the estimation of filters and receptor fields. There, a blind deconvolution algorithm was used to "dejitter" and realign

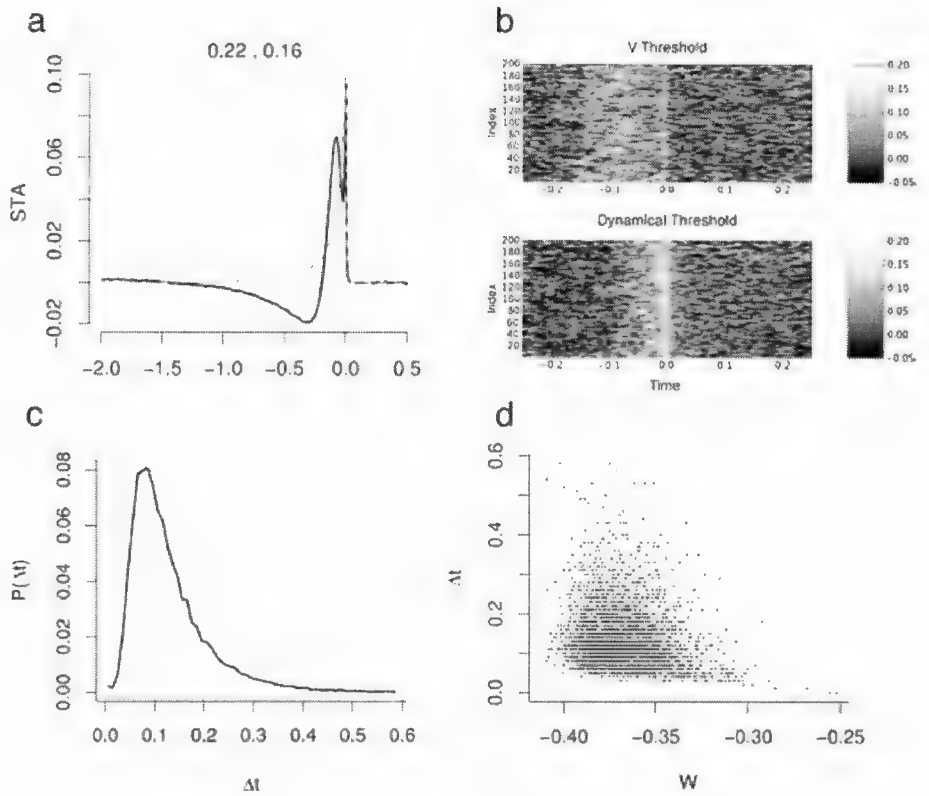


Figure 5: Effect of a dynamical threshold in the FN model. (a) STA of the full FN system with a constant V threshold. As before, a linear fit by the kernels is also shown. (b) Comparison of spike-triggered stimuli with different threshold choices in the FN model. (c) Temporal point-spread function induced by the choice of threshold. Δt is the time delay from crossing the dynamical threshold to the fixed V threshold. (d) Time delay of each spike-triggered stimulus plotted against the value of W at its threshold crossing point.

the spike-triggered stimuli, which dramatically sharpened the estimate of the spike-triggering stimulus. In a real system, jitter may indeed be due at least partly to noise, but it is also possible that a component of such jitter is deterministic and due to variability in the point of dynamical threshold crossing, as in the FN case. Blind deconvolution may then be viewed as an empirical approach to recovering a dynamical threshold based on spike times originally recorded using a voltage threshold. It is interesting to note that estimated jitter in such a case is correlated with projection onto the STA derivative.[6]

---

[6] This correlation arises because the STA derivative is the linear approximation to time translation of the STA. Hence, if a current history aligns with the STA better under a small time translation or jitter, it will have a projection onto the STA derivative proportional

4 Fully Nonlinear Systems

In this section, we discuss the covariance analysis of the full FN system and compare the result with the same analysis of the first- and second-order approximation.

We identified approximately 2×10^6 spikes first using a voltage threshold and then backtraced each trajectory to the point where it crosses the dynamical threshold shown in Figure 1. The trajectory might cross the threshold multiple times before spiking due to the noisy input. In this case, we used the first crossing point after the trajectory diverges from that of the second-order approximation. This is based on the assumption that the second order is a sufficiently good approximation of the system in the subthreshold regime. For both the first- and second-order system, crossing of the dynamical threshold is used to identify spikes. Right after a spike, we imposed a postspike inhibitory period equal to about a "spike width," as empirically determined from the full system.

The results are shown in Figures 6 and 7. Figure 6a shows that the STA of the first-order approximation is quite similar to that of the full system, and the second-order STA is even closer. From the covariance analysis, we see that the FN neuron, like the HH model (Agüera y Arcas et al., 2003), has one dominant negative and one dominant positive eigenvalue. However, the results from the covariance analysis of the three systems differ considerably. Both the second-order and the full system show a relatively large number of significant eigenvalues. This is due to the contribution of the second order $g_2[V]$ and $g_2[W]$ (and higher-order) kernels in equation 2.1. However, these are relatively suppressed compared to the first-order modes. Further, the spectrum derived from the second-order system, Figure 6b, has no positive significant eigenvalue.

Figure 6c provides more detailed information. The first-order case is just as we have seen previously, with three modes that are well described by linear combinations of the linear kernels. In the second-order case, v_1 and v_2 are comparable to those of the first-order. v_3, which is not approximated by the first-order kernels, must arise from the second-order kernels. In this regard, the full system is well matched with the second-order approximation. There also, v_3 is comparable to that in the second order, and $v_{1,2}$ are derived from the linear kernels. However, the full system also has a positive mode, v_+, from the linear kernels, which resembles the first order rather than the second.

---

to that jitter. In section 2, we demonstrated that the STA and its derivatives are, up to a linear transformation, the filters associated with the dynamical variables of the linearized system. Therefore, there is a relationship between blind deconvolution to optimize fit to an STA and estimation of a dynamical threshold based on the "hidden" (non-V) dynamical variables.

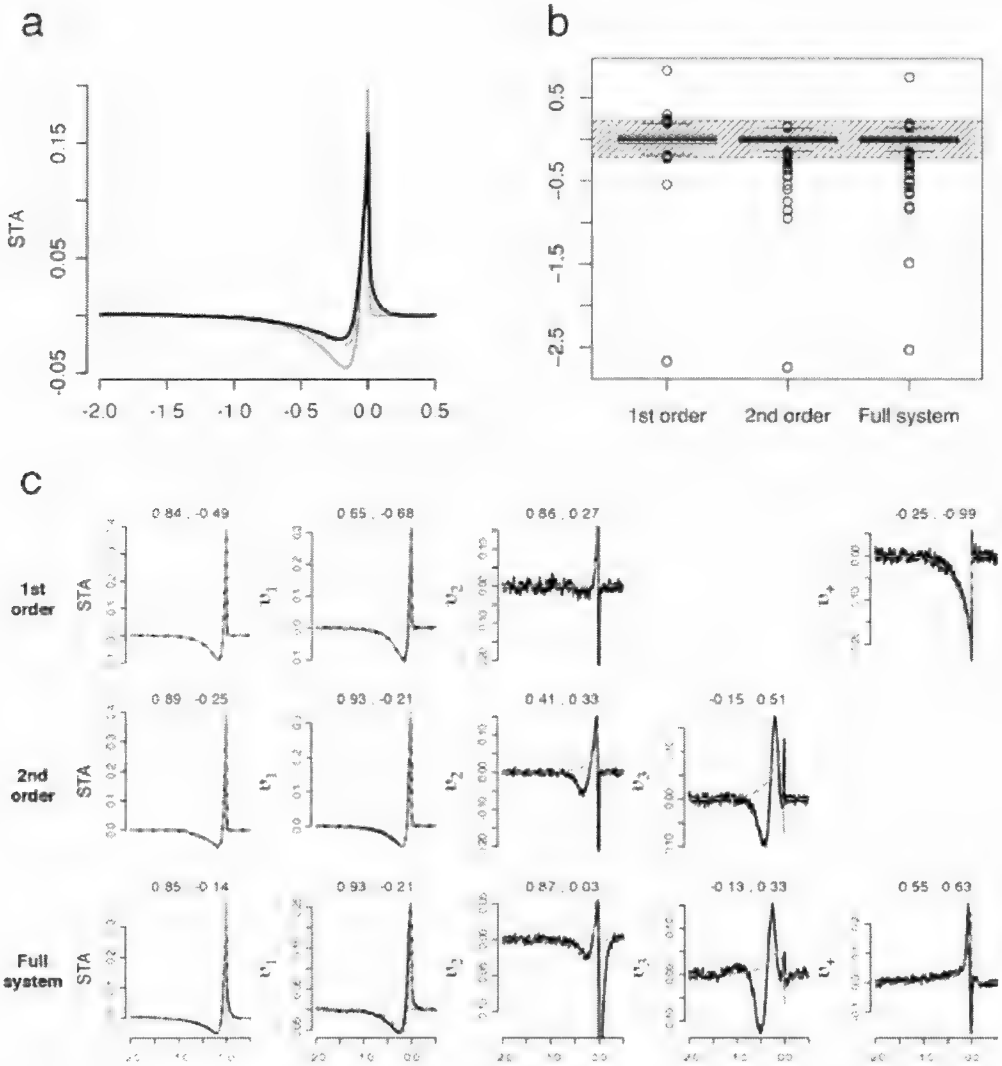


Figure 6: (a) STAs of an FN model neuron, and the first (gray) and second (dashed) order approximation. (b) Spectra of covariance matrices of the FN model and approximations. (c) STAs and covariance modes for the first- and second-order approximations and the full FN system.

The geometry of the spike-triggering stimulus projections is shown in Figure 7a. The first order, not surprisingly, recapitulates the curved threshold case in Figure 2d. However, the second order is more like the (V, W)-linear threshold case, while again the full system resembles the curve in Figure 2d. A possible explanation is that each system probes different parts of the dynamical threshold. Figure 7b marks the density of threshold crossing points for each model. In contrast to the first-order and the full system, which access a large section of the threshold with nontrivial curvature, the second order probes only a small and almost linear section. This may be the reason for the lack of a positive eigenvalue. As in the toy models

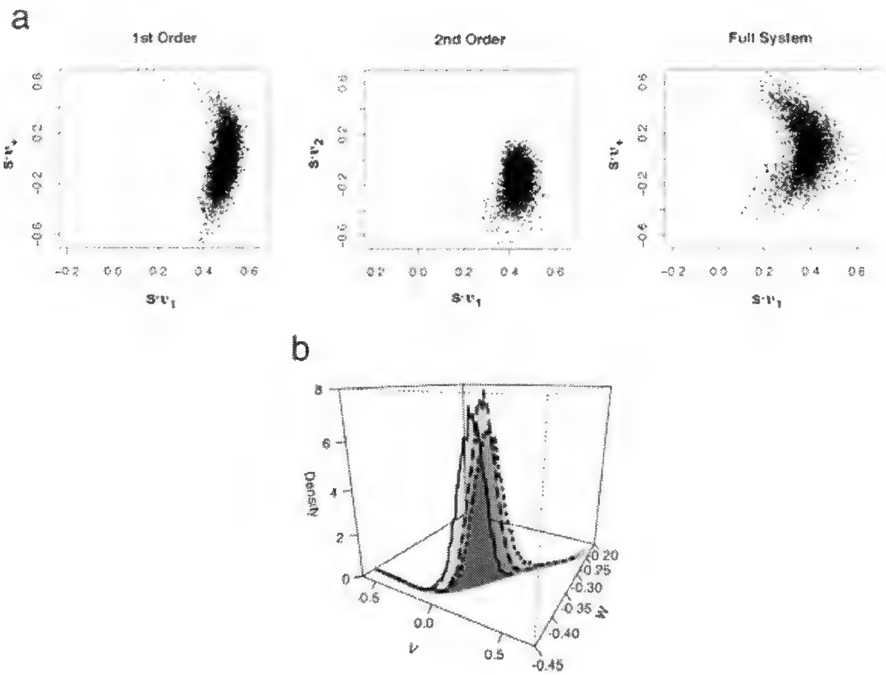


Figure 7: (a) Spike-triggered stimuli projected into a feature space defined by $v_{1,+}$ for the first-order case and full system and $v_{1,2}$ for the second-order system. (b) The density of threshold crossing points in the first-order (solid), second-order (dashed), and full (dotted) systems, plotted along the threshold curves in the $V - W$ plane.

in section 2, the contributions of the relevant dimensions are determined not only by local information (filters) but also by the global structure of a multidimensional or dynamical threshold.

5 Abbott-Kepler Model

In this section, we apply the same analysis to a two-dimensional model which is more nonlinear and more realistic than the FN model.

Abbott and Kepler (1990) developed a two-dimensional reduction of the Hodgkin-Huxley model, based on the observation that there is a separation of timescales between the faster m and the slower n and h variables. m is then replaced with its asymptotic value at the membrane voltage V, while n and h are controlled by another voltage variable U. The equations of the Abbott-Kepler (AK) model are of the form

$$C\frac{dV}{dt} = f(V, U) + I(t), \tag{5.1}$$

$$\frac{dU}{dt} = g(V, U), \tag{5.2}$$

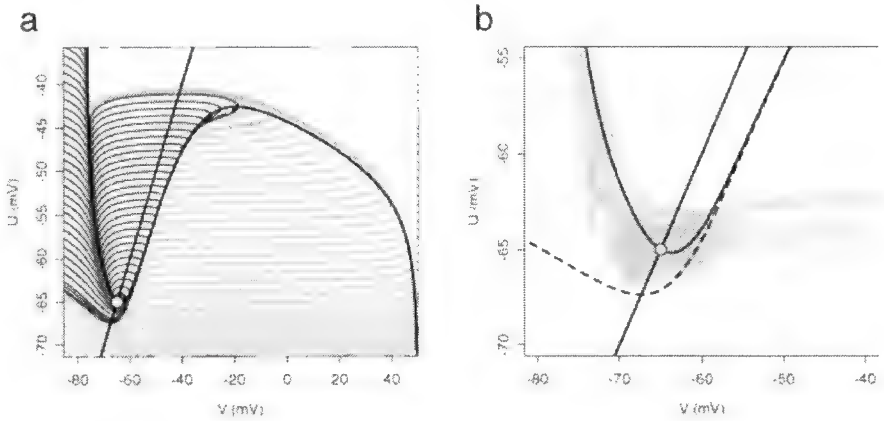


Figure 8: (a) Phase plane of the Abbott-Kepler model with trajectories, nullclines, and a threshold, in the absence of input. (b) AK phase plane with the injected noisy input.

where $f(V, U)$ and $g(V, U)$ are nonlinear functions in V and U. Their derivation is briefly sketched in appendix C, and we refer to the original paper (Abbott & Kepler, 1990) for further detail. The nullcline for U is given by $g(V, U) = 0$, which is satisfied by $V = U$. The V nullcline, $f(V, U) = 0$, is more complicated and obtained numerically. The two nullclines intersect at a fixed point $V = U = -65\,\text{mV}$.

Figure 8a shows the phase plane of this model with zero input current. Like Figure 1, the threshold structure, which can be obtained numerically, is visible. Due to the strong nonlinearity, spiking trajectories are well defined on the phase plane, and the threshold has less ambiguity than for the FN model. Again, we try to identify the dynamics of the system in the subthreshold regime with the first-order approximation. Unlike an FN neuron, the Jacobian of this system has complex eigenvalues $\lambda_{\pm} = -0.2118 \pm i0.4035\,\text{ms}^{-1}$, and therefore the first-order kernels $\epsilon^{(1)}_{V,U}$ oscillate, Figure 10a. This is consistent with the oscillatory linearized behavior associated with the full Hodgkin-Huxley model near equilibrium.

Before we carry out covariance analysis on this model, we examine the effect of a dynamical threshold, as in section 3. Figure 9a shows the spike-triggering stimuli aligned with both a fixed threshold in V, chosen as $V = -40\,\text{mV}$ to unambiguously select spiking trajectories, and the dynamical threshold. The typical time shift is of order of 1 ms or less; the overall STA suffers from only a slight time or phase shift when the V threshold is used. However, in the small timescale of 1 ms or less, the V-threshold STA differs from that in the dynamical threshold case, which is characterized by a large delta-function component at the spike time. Both STAs are well approximated by linear combinations of $\epsilon^{(1)}_{V,U}$, up to a small deformation around −15 ms due to the effects of multiple spikes.

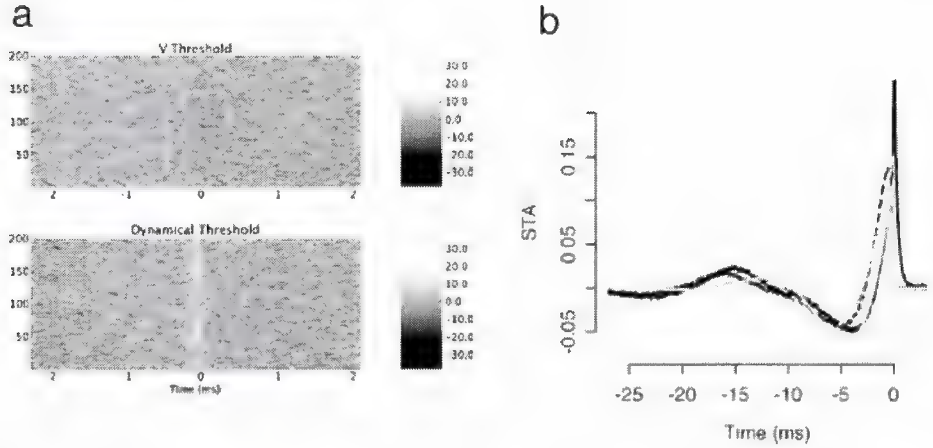


Figure 9: (a) Spike-triggering current histories with two different threshold schemes in the AK model. (b) STAs of the AK model with the fixed $V = -40\,\text{mV}$ threshold (dashed) and dynamical threshold (solid). Also, they are least squares fitted by linear combinations of $\epsilon_{V,U}$ (gray).

Figure 10 shows the results from covariance analysis carried out with approximately 2×10^6 spikes for gaussian white noise stimuli with various variances. For comparison with previous results, we selected three eigenmodes corresponding to the leading two negative and the largest positive eigenvalues. We find that they are reasonably well approximated by linear combinations of $\epsilon_{V,U}^{(1)}$ in most cases, although they are sometimes affected by a δ-function at $t = 0$ and a large multispike effect at high variances. The multispike effect can be eliminated by considering only the isolated spikes when the spike rate is low (Agüera y Arcas & Fairhall, 2003). At higher variances, isolated spikes are rare, and there is a stronger influence of oscillating "silence modes" (Agüera y Arcas et al., 2003). As in the FN neuron case, we identify modes other than those in Figure 10c as "nonlinear modes."

We compare results at different variances with our discussion in section 2.4. Some features of variance dependence are shared with the toy model in section 2.4: the eigenvalue spectrum drifts, and the corresponding modes rotate among themselves. However, the Abbott-Kepler modes also exhibit more complicated behavior. Figure 11 shows projections of spike-triggered stimuli onto two distinguished eigenvectors $v_{1,+}$ and corresponding threshold crossing points. At low variance, the system crosses mostly one side of the threshold while the projections trace out the curvature of the threshold segment. As the variance increases, some crossing points begin to appear on the left side of the threshold. However, this does not overcome the expansion of a crossing point distribution on the right side, and the modes corresponding to this direction dominate. At high variance, there are many crossing points on both sides, reflected in the bimodal distribution of the

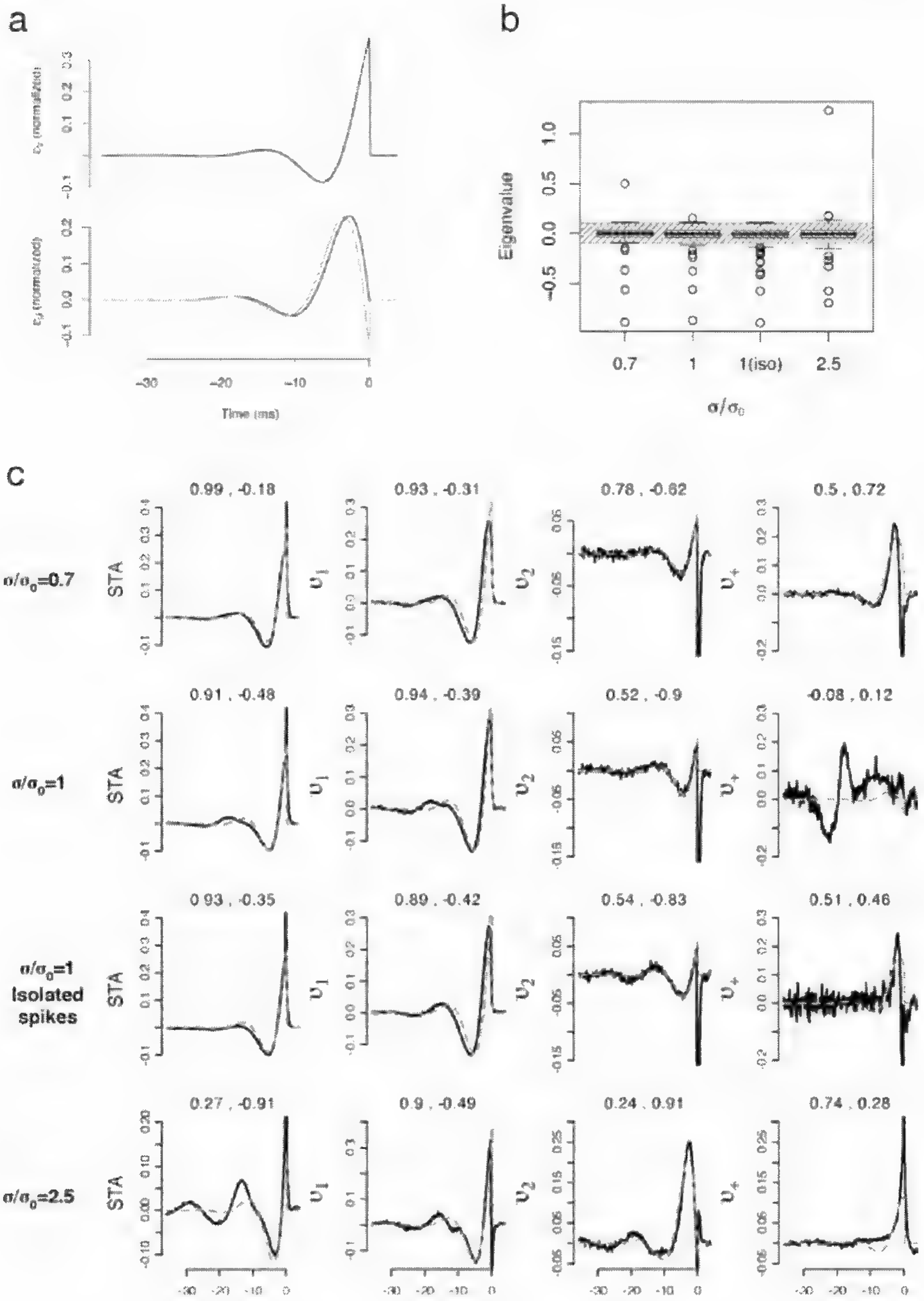


Figure 10: (a) Normalized first-order kernels $\bar{\epsilon}_{V,U}$ of the AK model and the component of ϵ_U orthogonal to ϵ_V, normalized (gray). (b) Eigenvalue spectrum of the covariance matrix of the AK model with each variance. $\sigma_0 = 13$pA. (c) Eigenmodes of the covariance matrices. As previously, $\upsilon_{1,2}$ and υ_+ correspond, respectively, to the two smallest negative eigenvalues and the largest positive one.

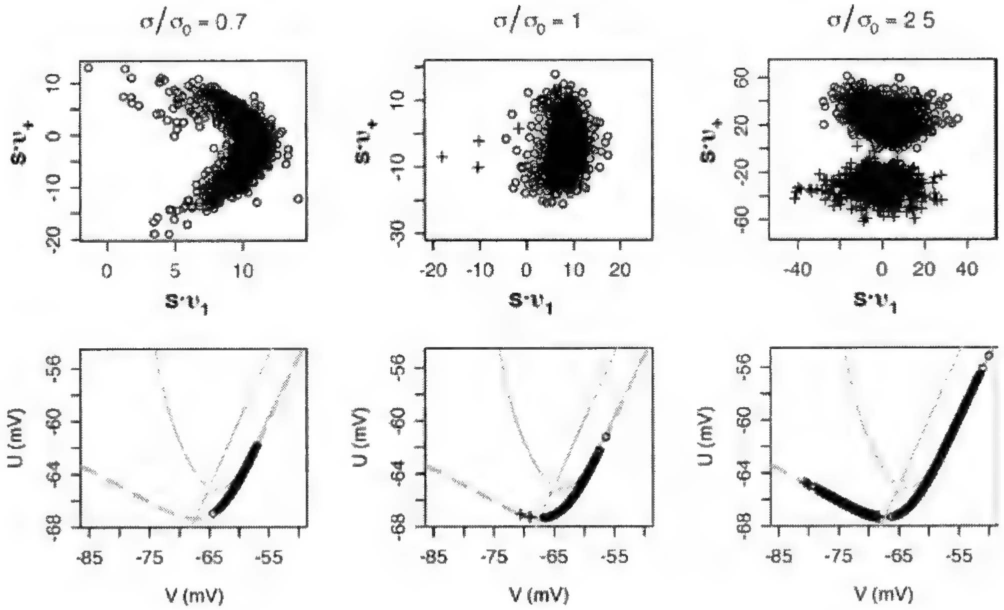


Figure 11: Projection of spike-triggering stimuli onto v_1 and v_+ in the AK model. Below are threshold crossing points in the phase plane of each case, marked by crosses, $V \leq -67\,\text{mV}$, and circles, when $V \geq -67\,\text{mV}$, over the dynamical threshold (dashed). The other gray curves are the nullclines, with their intersection, the fixed point, marked by a circle.

projections, previously seen in the toy models. Thus, this is another example of how the results of reverse correlation analysis are affected by both the filter properties and the interaction of the stimulus ensemble with the threshold geometry.

So far, we have discussed only two-dimensional dynamical models. We began with some artificial toy models with purely linear subthreshold dynamics and proceeded to the minimal FitzHugh-Nagumo spiking model. We applied the lessons learned there to the more nonlinear Abbott-Kepler model. We will conclude with a partial analysis of the higher-dimensional Hodgkin-Huxley model.

6 Hodgkin-Huxley Model

Higher-dimensional systems require nontrivial extensions of the methodology we have used with two-dimensional systems. First, it is much harder to find a dynamical threshold, which could now be a multidimensional hypersurface rather than a curve. If we do not align the spike-triggered stimuli according to the dynamical threshold, the obtained filters may include a distribution of time delays. For example, the reverse correlation analysis may result in broadened filters,

$$\tilde{f}_\mu(t) = \int d\tau\, f_\mu(\tau) p_\sigma(t-\tau), \tag{6.1}$$

where $p_\sigma(\tau)$ is a point-spread function, depending on the input variance σ^2, as in Figure 5c.[7]

However, we can still gain some insights from $\tilde{f}_\mu$. If $p(t)$ has a narrow support, say less than a millisecond, it has negligible effects for larger time windows. It is also possible that some consequences of our analysis will still apply to $\tilde{f}_\mu$ with only a few reasonable assumptions. For example, if p_σ is of limited temporal extent, then the derivatives of $\tilde{f}_\mu$ can be simply

$$\tilde{f}_\mu^{(n)} \approx \int d\tau \, f_\mu^{(n)}(\tau) p_\sigma(t-\tau). \tag{6.2}$$

Therefore, as we discussed in section 2, the linear modes among $\{\tilde{f}_\mu\}$ should still be (approximately) closed under time differentiation.

To demonstrate this, we use the four-dimensional Hodgkin-Huxley model, and the following strategy: we select the linear modes out of the significant filters from covariance analysis. Now, since the STA is approximately their linear combination, the time derivatives of the STA should be written in terms of the linear combinations as well if they are closed under time differentiation. Figure 12 shows that this holds. In the low-variance case, Figure 12a, the three linear modes chosen provide good fits to the time derivatives of the STA. The high-variance case, Figure 12b, is affected by multispike effects and a filtering artifact, but it is also well fitted. Therefore, we can conclude that the time derivatives of the STA span the same space as three linear modes. One might expect a fourth linear mode due to the dimensionality of the model, but this is not as significant as the others; this agrees with previous covariance analysis that the HH model can be well described as a quasi-three-dimensional system (Agüera y Arcas et al., 2003).

We note that it is not clear whether the linear modes in the two cases span the same feature space. This is difficult to ascertain since the point-spread function can in principle depend on the stimulus variance. In Figure 12c, we see that while v_1 of the high-variance case can be fitted by the linear modes of the low variance, the other modes show a small deviation even around 5 ms. This might indicate an interesting variance dependence as in section 2.4, but we will not pursue this issue in this article.

7 Summary

Here we have investigated the meaning or interpretation of the features derived from the spike-triggered covariance method applied to two simple

---

[7] In fact, it can be much more complicated than this. Since the time delay depends on the location of threshold crossing in a phase space, additional temporal correlations are introduced. Therefore, each filter may have a different point-spread function, or filters may depend on the time delay structure in a complicated way. Here will discuss only the simplest case.

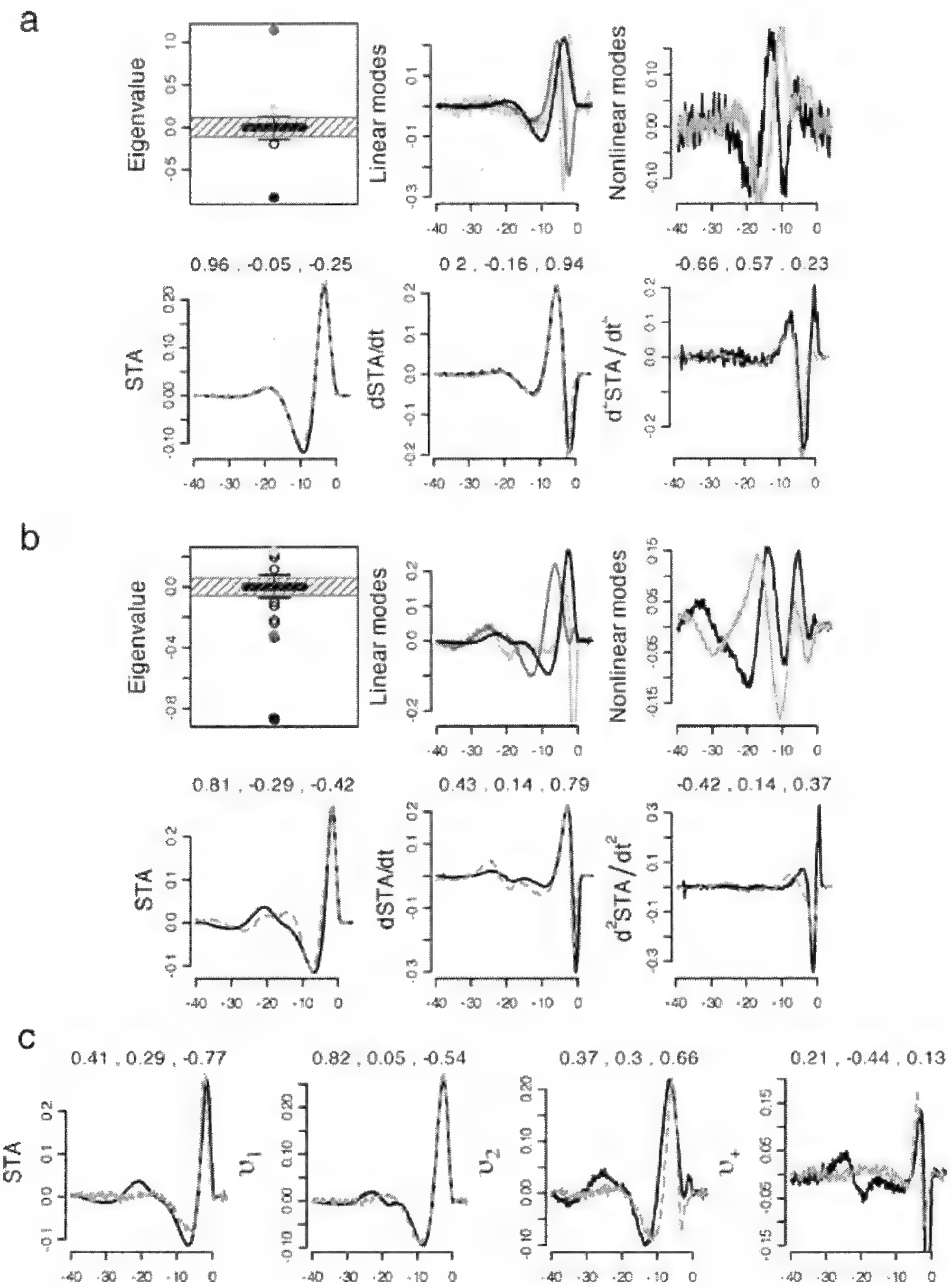


Figure 12: Covariance bases of the linear feature space of the Hodgkin-Huxley model recovered at (a) low and (b) high variance. The three most significant eigenvalues are denoted with the same shades as the corresponding linear eigenmodes. (c) The STA and linear modes of the high-variance case and their fits using the linear modes of the low-variance case in the HH model.

neuron models: the FitzHugh-Nagumo model, the minimal spiking neuron model, and the Abbott-Kepler model, a more faithful two-dimensional reduction of the Hodgkin-Huxley system. The power of white noise analysis is that it provides a data-driven method to reduce a high-dimensional dynamical system to a functional model that captures the essential computation of the system in terms familiar to systems neuroscience: a receptive field that filters the stimulus and a threshold function over the filtered stimulus. Our goal here was to analyze the output of a white noise analysis in terms of what it can reveal and how it depends on the underlying dynamical system. In this, our approach is distinct from the elegant work of Huys, Ahrens, & Paninski (2006), where responses to white noise stimuli are used to fit the parameters of a conductance-based model.

Dealing with simple two-dimensional systems, our observations of spiking dynamics in the phase plane motivated the following reduced model: dynamics in the phase plane are approximated by the perturbative expansion, in particular the linear approximation, and the system's nonlinearity is captured by a spiking threshold, determined for zero input, that extends through multiple dimensions. This model is a generalization of a basic filter-and-fire model to multiple dimensions and extends it in two significant ways. One is the identification of the filters with the dynamical variables of the system; the other is the generalization of the concept of the threshold.

The simplified model with linear dynamics and a multidimensional curved threshold was treated both analytically and phenomenologically, using numerical simulation and reverse correlation (covariance) analysis of spike-triggered stimuli. This led to several insights, some of which are related to existing observations. First, for this case, the feature space derived from covariance analysis is spanned by the linear kernels. Since the set of kernels is closed under time differentiation, the same feature space can also be spanned by a generic linear combination of the kernels, such as the spike-triggered average and its n time derivatives. However, not every kernel contributes as a relevant feature. The threshold structure plays a role of selecting the relevant ones from the kernels: for example, a linear threshold selects a single filter and its time derivative. In general, a threshold spanning a d-dimensional subspace will select $d + 1$ features from among linear combinations of the kernels. We show further for this model that the distribution of spike-triggered stimuli in the covariance feature space corresponds to that of threshold crossing points, up to a suitable linear transformation.

We also used this model to illustrate the effects of the interaction of the threshold nonlinearity with the stimulus ensemble. We show that a complex threshold geometry leads to a nontrivial variance dependence of the eigenmodes and eigenvalues of covariance analysis and, even more so, the spike-triggered average. In the linearized subthreshold model, the subspace of eigenmodes is not changed, but the spike-triggering ensemble may rotate through this subspace, leading to a variance-dependent spike-triggered average. This is one example of stimulus-variance dependence in a nonlinear, nonadapting system (Yu & Lee, 2003; Borst, Flanagin, & Sompolinsky, 2005). We show this effect in the analysis of the Abbott-Kepler model neuron.

The identification of a curved or dynamical threshold raises an issue in reverse correlation analysis regarding the determination of spike time. For the FitzHugh-Nagumo and Abbott-Kepler models, we compared two cases: when spike times are identified using a threshold in voltage and when the spike times are found using the crossing of an identified curve that leads to spiking for zero current input. Surprisingly, the spike-triggered averages showed remarkable differences. For a voltage threshold, the time delay from the earlier threshold crossing point blurs the estimated filters. It is shown that using the dynamical threshold leads to a better fitting of the estimated filters by the system's linear kernels, as well as improving the sharpness and substantially altering the shape of the STA.

In all of our discussion here, we have concentrated on the subthreshold dynamics bringing the system to the point of threshold. We do not analyze, and our simplified models do not accommodate, the dynamics of the system immediately after a spike. In the phase-plane picture, spiking reinjects the system into the subthreshold regime in a nonrandom way, affecting the subsequent probability flux to threshold, analogous to the one-dimensional reset of the integrate-and-fire neuron treated in Paninski, Lau, & Reyes, (2003). In our approach, we have assumed that the system has remained below threshold for long enough that its location in the subthreshold space has been randomized by the driving current. This corresponds to an analysis of isolated spikes only, a simplification we have used before (Agüera y Arcas et al., 2003) and employed here for the covariance analysis. A number of works have treated this issue explicitly with a variety of methods: treating the interspike interval as the primary symbol (de Ruyter van Steveninck & Bialek, 1988; Rieke et al., 1997), solving for the interspike interaction using the interspike intervals (Pillow & Simoncelli, 2003), simultaneously solving for the linear kernel over stimulus history and spike history using the autoregressive moving average (Powers et al., 2005; Truccolo, Eden, Fellows, Donoghue, & Brown, 2005), and fitting parameters of an explicit model for the effective postspike current (Kistler et al., 1997; Keat et al., 2001; Paninski et al., 2004; Pillow et al. 2005). While the simplification we have used here allows us to find direct connections between the covariance modes without the confound of the interspike interaction (Agüera y Arcas et al., 2003; Agüera y Arcas, & Fairhall, 2003) it is clearly not a complete model for spiking responses. A first step toward a more complete spiking model may be to consider the perturbed subthreshold distributions induced by the influx of trajectories following spikes. This is in effect a mean field approximation, taking into account the overall spike rate for a given stimulus ensemble. Further steps could be taken by introducing a return map deterministically relating the point of threshold crossing to a point of return into the subthreshold domain. The integrate-and-fire model is the most trivial implementation of such a return map; an equivalent map in multiple dimensions would reintroduce all spike trajectories into the subthreshold domain at an identified point (this is $V = 0$ for integrate-and-fire). Such a many-to-one map implies that the neuron's state is completely reset by a spike, which is incorrect for neurons with slow conductances that modulate spike afterpotentials. A less degenerate map seems more appropriate for such cases. Another important step is the addition of these slow conductances. Using our formalism, such conductances may be representable simply as additional dimensions of threshold curvature with corresponding longer-timescale stimulus filters.

White noise analysis allows the derivation of intuitive functional models for neural computation: What does the neuron compute? In this article, we have drawn concrete correspondences between the components of these functional models and parameters of the underlying dynamical system.

Appendix A: Volterra Expansion of a Dynamical System

Let us consider an n-dimensional dynamical system, perturbed by an external input $I(t)$ as equation 2.1. For convenience, we introduce a dimensionless expansion parameter η, with which we will expand the equation. More precisely, equation 2.1 becomes

$$\partial_t y_k = f_k(y_1, y_2, \ldots, y_N) + \eta \delta_{k0} I(t). \tag{A.1}$$

We also have the perturbation expansion of y_k as

$$y_k = y_k^{(0)} + \eta y_k^{(1)} + \eta^2 y_k^{(2)} + \cdots. \tag{A.2}$$

By plugging equation A.2 into equation A.1, we obtain

$$\begin{aligned}\partial_t \left(\eta y_k^{(1)} + \eta^2 y_k^{(2)} + \cdots\right) = \eta I(t)\delta_{k0} &+ \frac{\partial f_k(y^{(0)})}{\partial y_m} \eta y_m^{(1)} \\ &+ \frac{1}{2!} \frac{\partial^2 f_k(y^{(0)})}{\partial y_m \partial y_n} \eta^2 y_m^{(1)} y_n^{(1)} \\ &+ \frac{\partial f_k(y^{(0)})}{\partial y_m} \eta^2 y_m^{(2)} + \cdots.\end{aligned}$$

By comparing two sides order by order, we obtain a series of equations satisfied by the perturbative expansion at each order as

$$\mathcal{D}_{km} y_m^{(1)} = I\,\delta_{k0}, \quad \mathcal{D}_{km} = \delta_{km}\partial_t - \left.\frac{\partial f_k(x^{(0)})}{\partial x_m}\right|_{y^{(0)}},$$

which is equation 2.2,

$$\mathcal{D}_{km} y_m^{(2)} = \frac{1}{2!} H_{kmn} y_m^{(1)} y_n^{(1)}, \quad H_{kmn} = \frac{\partial^2 f_k(y^{(0)})}{\partial y_m \partial y_n}, \tag{A.3}$$

which is equation 2.9, and so on.

These equations can be solved recursively by using a kernel $\epsilon^{(1)} = \mathcal{D}^{-1}$, which therefore satisfies

$$\mathcal{D}_{km}\epsilon_{ml}^{(1)}(t, t') = \delta_{kl}\delta(t - t').$$

Now the solution of the first order can be written as

$$y_k^{(1)}(t) = \int dt'\, \epsilon_k^{(1)}(t - t') I(t'),$$

where $\epsilon_k^{(1)} = \epsilon_{k0}^{(1)}$. Also from equation A.3,

$$\begin{aligned} y_k^{(2)} &= \frac{1}{2!}\int dt'\, \epsilon_{kl}^{(1)}(t - t') \\ &\quad \times \left(H_{lmn}\int ds_1 ds_2\, \epsilon_m^{(1)}(t' - s_1)\epsilon_n^{(1)}(t' - s_2) I(s_1) I(s_2)\right) \\ &= \frac{1}{2!}\int ds_1 ds_2\, \epsilon_k^{(2)}(t - s_1, t - s_2) I(s_1) I(s_2), \end{aligned}$$

$$\epsilon_k^{(2)}(s_1, s_2) = \int dt'\, H_{lmn}\epsilon_{kl}^{(1)}(t')\epsilon_m^{(1)}(s_1 - t')\epsilon_n^{(1)}(s_2 - t'),$$

which is equation 2.10. Since this procedure can be carried out to higher orders, the higher-order kernels are outer products of the linear kernels.

In addition, the first-order kernels are related to each other by simple time differentiation. To show this, we rewrite equation 2.4 as follows:

$$\epsilon_m^{(1)}(t - t') = \mathcal{D}_{mn}^{-1}\delta_{n0}\delta(t - t'). \tag{A.4}$$

Now the derivatives can be computed as

$$\begin{aligned}\partial_t \epsilon^{(1)}_m &= \partial_t \mathcal{D}^{-1}_{mn}\delta_{n0}\delta(t-t')\\ &= (\partial_t \delta_{ml} - J_{ml} + J_{ml})\mathcal{D}^{-1}_{ln}\delta_{n0}\delta(t-t')\\ &= \mathcal{D}_{ml}\mathcal{D}^{-1}_{ln}\delta_{n0}\delta(t-t') + J_{ml}\epsilon^{(1)}_l\\ &= \delta_{m0}\delta(t-t') + J_{ml}\epsilon^{(1)}_l.\end{aligned} \tag{A.5}$$

This shows that the nonsingular part of the first-order kernels can be obtained by the linear combination of the derivatives of other kernels.

Appendix B: A Linear Model from the FitzHugh-Nagumo System

Here we derive the linearization of the FitzHugh-Nagumo model, beginning with equations 1.5 and 1.6. The fixed point can be obtained by simultaneously solving the nullcline equations. We denote it by (V_0, W_0) and expand the system around this point.

First, the Jacobian is

$$J = \begin{pmatrix} F'(V_0)/\psi & -1/\psi \\ 1 & -b \end{pmatrix},$$

where $F(V) = V(1-V)(a+V)$, and this defines a linear system:

$$\begin{aligned}\partial_t V &= \frac{F'(V_0)}{\psi}V - \frac{1}{\psi}W + I(t)\\ \partial_t W &= V - bW.\end{aligned}$$

J has eigenvalues $\lambda_\pm$:

$$\begin{aligned}\lambda_\pm &= \frac{1}{\psi}\left[f_-(V_0) \pm \sqrt{f_+(V_0)^2 - \psi}\right]\\ &= -b + \left[f_+(V_0) \pm \sqrt{f_+(V_0)^2 - \psi}\right]\\ &= -b + \kappa_\pm,\end{aligned} \tag{B.1}$$

where $f_\pm(V) = (F'(V) \pm b\psi)/2$. As in equation 2.6, J is diagonalized by a matrix U,

$$U = \frac{1}{\lambda_+ - \lambda_-}\begin{pmatrix} 1 & \lambda_- + b \\ -1 & -(\lambda_+ + b) \end{pmatrix}.$$

From equation 2.8, we obtain the first-order kernels, which solve the linear system

$$\begin{aligned}\epsilon_V^{(1)}(t) &= e^{-bt}\partial_t S(t)H(t),\\ \epsilon_W^{(1)}(t) &= e^{-bt}S(t)H(t),\end{aligned} \tag{B.2}$$

where $S(t) = (e^{\kappa_+ t} - e^{\kappa_- t})/(\kappa_+ - \kappa_-)$. $\epsilon_{V,W}^{(1)}(t)$ with our choice of parameters is drawn in Figure 2a.

Appendix C: Derivation of the Abbott-Kepler Model

In this section, we show the derivation of a two-dimensional neuron model used in section 5. For further details, we refer to the original paper (Abbott & Kepler, 1990).

We begin with a Hodgkin-Huxley equation, which is defined by equation 1.1 and the following parameters:

$$g_L = \bar{g}_L, \qquad g_K = \bar{g}_K n^4, \qquad g_{Na} = \bar{g}_{Na} m^3 h. \tag{C.1}$$

$$\begin{aligned}\tau_z(V)\frac{dz}{dt} &= \bar{z}(V) - z, \qquad \tau_z = \frac{1}{\alpha_z + \beta_z},\\ \bar{z} &= \frac{\alpha_z}{\alpha_z + \beta_z}, \qquad z = m, n, h\end{aligned} \tag{C.2}$$

$$\begin{aligned}\alpha_m &= \frac{.1(V+40)}{1-\exp[-.1(V+40)]}, \qquad \beta_m = 4\exp[-.0556(V+65)],\\ \alpha_h &= .07\exp[-.05(V+65)], \qquad \beta_h = \frac{1}{1+\exp[-.1(V+35)]},\\ \alpha_n &= \frac{.01(V+55)}{1-\exp[-.1(V+55)]}, \qquad \beta_n = .125\exp[-.0125(V+65)].\end{aligned} \tag{C.3}$$

The key observation is that τ_m is much smaller than $\tau_{n,h}$, while τ_h and τ_n are mutually comparable. Therefore, m can be approximated by its value at equilibrium, $\bar{m}(V)$, and h and n can be represented by the same equilibrium voltage, U. In other words,

$$m \approx \bar{m}(V), \qquad h \approx \bar{h}(U), \qquad n \approx \bar{n}(U).$$

Using this, we obtain equation 5.1 as

$$\begin{aligned}F &= \sum_{i=L,K,Na} g_i(V - E_i)\\ &\approx \bar{g}_L(V - E_L) + \bar{g}_K \bar{n}(U)^4(V - E_K) + \bar{g}_{Na}\bar{m}(V)^3\bar{h}(U)(V - E_{Na})\\ &= -f(V, U).\end{aligned}$$

An equation for U, equation 5.2, is obtained by requiring that time dependence of the active current F due to h and n in the Hodgkin-Huxley model is mimicked by $\bar{h}(U)$ and $\bar{n}(U)$. This implies

$$\frac{dF}{dh}\frac{dh}{dt} + \frac{dF}{dn}\frac{dn}{dt} \approx \left(\frac{\partial f}{\partial \bar{h}}\frac{d\bar{h}}{dU} + \frac{\partial f}{\partial \bar{n}}\frac{d\bar{n}}{dU}\right)\frac{dU}{dt}. \tag{C.4}$$

Again, we approximate equation C.2 in the same way,

$$\frac{dz}{dt} \approx \frac{1}{\tau_z(V)}\left(\bar{z}(V) - \bar{z}(U)\right), \qquad z = h, n.$$

Plugging this in equation C.4, we can solve for dU/dt as a function of V and U, which we denoted by $g(V, U)$ in equation 5.2 as

$$g(V, U) = \frac{\bar{g}_{Na}(V - E_{Na})\bar{m}(V)^3(\bar{h}(V) - \bar{h}(U))/\tau_h(V) + 4\bar{g}_K(V - E_K)\bar{n}(U)^3(\bar{n}(V) - \bar{n}(U))/\tau_n(V)}{\bar{g}_{Na}(V - E_{Na})\bar{m}(V)^3\bar{h}'(U) + 4\bar{g}_K(V - E_K)\bar{n}(U)^3\bar{n}'(U)}.$$

References

Abbott, L. F., & Kepler, T. (1990). Model neurons: From Hodgkin-Huxley to Hopfield. In L. Garrido (Ed.), *Statistical mechanics of neural networks* (pp. 5–18). Berlin: Springer-Verlag.

Agüera y Arcas, B., & Fairhall, A. (2003). What causes a neuron to spike? *Neural Computation, 15*, 1715–1749.

Agüera y Arcas, B., Fairhall, A., & Bialek, W. (2003). Computation in a single neuron: Hodgkin and Huxley revisited. *Neural Computation, 15*, 1789–1807.

Aldworth, Z., Miller, J., Gedeon, T., Cummins, G., & Dimitrov, A. (2005). Dejittered spike-conditioned stimulus waveforms yield improved estimates of neuronal feature selectivity and spike timing precision of sensory interneurons. *J. Neurosci., 25*, 5323–5332.

Azouz, R., & Gray, C. (2000). Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons in vivo. *Proc. Natl. Acad. Sci. USA, 97*, 8110–8115.

Berry II, M. J., & Meister, M. (1999). The neural code of the retina. *Neuron, 22*, 435–450.

Bialek, W., & de Ruyter van Steveninck, R. R. (2005). *Features and dimensions: Motion estimation in fly vision.* Preprint, arXiv:q-bio/0505003 [q-bio.NC].

Borst, A., Flanagin, V. L., & Sompolinsky, H. (2005). Adaptation without parameter change: Dynamic gain control in motion detection. *Proc. Natl. Acad. Sci. USA, 102*(17), 6172–6176.

Brenner, N., Bialek, W., & de Ruyter van Steveninck, R. R. (2000). Adaptive rescaling maximizes information transmission. *Neuron, 26*, 695–702.

de Ruyter van Steveninck, R. R., & Bialek, W. (1988). Real-time performance of a movement sensitive in the blowfly visual system: Information transfer in short spike sequences. *Proc. Roy. Soc. Lond. B, 234*, 379–414.

DeWeese, M. (1995). *Optimization principles for the neural code.* Unpublished doctoral dissertation, Princeton University.

Fairhall, A. L., Burlingame, C. A., Narasimhan, R., Harris, R. A., Puchalla, J. L., & Berry II, M. J. (2006). Selectivity for multiple stimulus features in retinal ganglion cells. *J. Neurophysiol., 96*, 2724–2738.

FitzHugh, R. (1961). Impulse & physiological states in models of nerve membrane. *Biophysics J., 1*, 445–466.

Gerstner, W., & Kistler, W. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge: Cambridge Press.

Gradshteyn, I. S., & Ryzhik, I. M. (2000). *Table of integrals, series, and products* (6th ed.). San Diego: Academic Press.

Horwitz, G., Chichilnisky, E., & Albright, T. (2005). Blue-yellow signals are enhanced by spatiotemporal luminance contrast in macaque V1. *J. Neurophysiol., 93*(4), 2263–2278.

Hutcheon, B., & Yarom, Y. (2000). Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends Neurosci., 23*(5), 216–222.

Huys, Q. J., Ahrens, M. B., & Paninski, L. (2006). Efficient estimation of detailed single-neuron models. *J Neurophysiol., 96*, 872–890.

Izhikevich, E. (2001). Resonate-and-fire neurons. *Neural Networks, 14*, 883–894.

Izhikevich, E. M. (2006). *Dynamical systems in neuroscience: The geometry of excitability and bursting.* Cambridge, MA: MIT Press.

Keat, J., Reinagel, P., Reid, R. C., & Meister, M. (2001). Predicting every spike: A model for the responses of visual neurons. *Neuron, 30(3)*, 803–817.

Kistler, W., Gerstner, W., & van Hemmen, J. L. (1997). Reduction of the Hodgkin-Huxley equations to a single-variable threshold model. *Neural Computation, 9*, 1015–1045.

Maravall, M., Petersen, R. S., Fairhall, A. L., Arabzadeh, E., & Diamond, M. E. (2007). Shift in coding properties and maintenance of information transmission during adaptation in barrel cortex. *PLoS Biol., 5*, e19.

Marmarelis, P. Z., & Marmarelis, V. Z. (1974). Measurement of the Wiener kernels of a non-linear system by cross-correlation. *International Journal of Control, 2*, 234–254.

Marmarelis, V. (2004). *Nonlinear dynamic modeling of physiological systems.* Piscataway, NJ: IEEE Press.

Mehta, M. L. (1991). *Random matrices* (3rd ed.). New York: Academic Press.

Nagumo, J., Arimoto, S., & Yoshikawa, Z. (1962). An active pulse transmission line simulating nerve axon. *Proc. IRE, 50*, 2061–2071.

Paninski, L., Lau, B., & Reyes, A. (2003). Noise-driven adaptation: in vitro and mathematical analysis. *Neurocomputing, 52*, 877–883.

Paninski, L., Pillow, J., & Simoncelli, E. (2004). Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model. *Neural Comp., 16*, 2533–2561.

Petersen, R. (2003). Coding of dynamical whisker stimulation in the rat somatosensory cortex: A spike-triggered covariance analysis. *Society for Neuroscience*, 57.1.

Pillow, J. W., Paninski, L., Uzzell, V. J., Simoncelli, E. P., & Chichilnisky, E. J. (2005). Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. *J. Neurosci., 25*(47), 11003–11013.

Pillow, J. W., & Simoncelli, E. P. (2003). Biases in white noise analysis due to non-Poisson spike generation. *Neurocomputing, 52–54*, 109–115.

Powers, R., Dai, Y., Bell, B., Percival, D., & Binder, M. (2005). Contributions of the input signal and prior activation history on the discharge behavior of rat motoneurones. *J. Physiol., 562*, 707–724.

Rieke, F. (1991). *Physical principles underlying sensory processing and computation*. Unpublished doctoral dissertation, University of California, Berkeley.

Rieke, F., Warland, D., Bialek, W., & de Ruyter van Steveninck, R. R. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Rust, N., Schwartz, O., Movshon, J., & Simoncelli, E. (2005). Spatiotemporal elements of macaque V1 receptive fields. *Neuron, 46*, 945–956.

Simoncelli, E. P., Pillow, J., Paninski, L., & Schwartz, O. (2004). Characterization of neural responses with stochastic stimuli. In M. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.

Slee, S. J., Higgs, M. H., Fairhall, A. L., & Spain, W. J. (2005). Two-dimensional time coding in the auditory brainstem. *J. Neurosci., 25*(43), 9978–9988.

Touryan, J., Lau, B., & Dan, Y. (2002). Isolation of relevant visual features from random stimuli for cortical complex cells. *J. Neurosci., 22*, 10811–10818.

Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., & Brown, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J. Neurophysiol., 93*(2), 1074–1089.

Victor, J. D., & Shapley, R. M. (1979). The nonlinear pathway of Y ganglion cells in the cat retina. *J. Gen. Physiol., 74*(6), 671–689.

Victor, J., & Shapley, R. (1980). A method of nonlinear analysis in the frequency domain. *Biophys. J., 29*(3), 459–483.

Volterra, V. (1930). *Theory of functions and of integral and integrodifferential equations*. New York: Dover.

Wiener, N. (1958). *Nonlinear problems in random theory*. Cambridge, MA: MIT Press.

Yu, Y., & Lee, T. S. (2003). Dynamical mechanisms underlying contrast gain control in single neurons. *Phys. Rev. E, 68*(1 Pt. 1), 011901–011901.

Received May 31, 2006; accepted November 10, 2006.

J Comput Neurosci
DOI 10.1007/s10827-009-0142-x

Sensitivity of firing rate to input fluctuations depends on time scale separation between fast and slow variables in single neurons

**Brian Nils Lundstrom · Michael Famulare ·
Larry B. Sorensen · William J. Spain ·
Adrienne L. Fairhall**

Received: 20 August 2008 / Revised: 11 December 2008 / Accepted: 6 February 2009
© Springer Science + Business Media, LLC 2009

Abstract Neuronal responses are often characterized by the firing rate as a function of the stimulus mean, or the *f–I* curve. We introduce a novel classification of neurons into Types A, B−, and B+ according to how *f–I* curves are modulated by input fluctuations. In Type A neurons, the *f–I* curves display little sensitivity to input fluctuations when the mean current is large. In contrast, Type B neurons display sensitivity to fluctuations throughout the entire range of input means. Type B− neurons do not fire repetitively for any constant input, whereas Type B+ neurons do. We show that Type B+ behavior results from a separation of time scales between a slow and fast variable. A voltage-dependent time constant for the recovery variable can facilitate sensitivity to input fluctuations. Type B+ firing rates can be approximated using a simple "energy barrier" model.

Keywords Noise · Gain · *f–I* curve · Stimulus fluctuations · Single neuron · Time scales · Dynamical systems · Phase portrait · Hodgkin-Huxley · Slow adaptation · Slow AHP

Action Editor: Bard Ermentrout

B. N. Lundstrom (✉) · W. J. Spain · A. L. Fairhall
Department of Physiology and Biophysics,
University of Washington,
Seattle, WA 98195, USA
e-mail: lundbr@u.washington.edu

M. Famulare · L. B. Sorensen
Department of Physics, University of Washington,
Seattle, WA 98195, USA

W. J. Spain
Neurology Section,
Veterans Affairs Puget Sound Health Care System,
1660 South Columbian Way,
Seattle, WA 98108, USA

1 Introduction

The output of a neuron can be described in terms of precise spike timing or the mean firing rate (Shadlen and Newsome 1994; Konig et al. 1996; VanRullen et al. 2005). Both of these coding strategies convey information, possibly distinct, about the stimulus (Fairhall et al. 2001; Lundstrom and Fairhall 2006). Here, we focus on how the mean firing rate of a single-compartment neuron encodes the statistical properties of current inputs, which we approximate as filtered Gaussian noise (Destexhe et al. 2001, 2003; Moreno et al. 2002; for discussion, see Rauch et al. 2003; Richardson 2004; Rudolph and Destexhe 2005, 2006).

While firing rate generally increases with stimulus mean, the slope of the firing rate to input current (*f–I*) curve can be altered by the variance of the fluctuating input. This can be seen as a type of gain control (Chance et al. 2002; Fellous et al. 2003; Rauch et al. 2003; Higgs et al. 2006; Arsiero et al. 2007; Lundstrom et al. 2008), whereby the strength of background fluctuations modulates the response to the mean. More generally, the firing rate varies with, and thereby encodes, both mean and variance, and the sensitivity to one or the other can be interpreted as conferring on the neuron integrator-like or differentiator-like properties, respectively (Higgs et al. 2006; Lundstrom et al. 2008). For example, the average firing rate of layer 5 pyramidal neurons of the sensorimotor cortex recorded *in vitro* varies with input mean but depends less strongly on input variance (Chance et al. 2002; Rauch et al. 2003); these neurons can be thought of as integrators. The firing rates of other neurons, e.g. from the rat prefrontal cortex, are more sensitive to changes of input variance (Fellous et al. 2003; Arsiero et al. 2007) and can be classified as differentiators, or coincidence detectors, where this sensitivity to transients

Published online: 08 April 2009

can be facilitated by mechanisms underlying spike rate adaptation (Benda et al. 2005; Higgs et al. 2006; Prescott et al. 2006). However, a strict dichotomy between integrator and differentiator is artificial. Plotting *f–I* curves for a number of examples of recent *in vitro* data [Fig. 1(a)], it is clear that the firing rate is often sensitive to both input mean and variance, and that sensitivity varies over the neuron's dynamic range.

To what extent is this range of behavior captured by standard single-compartment model neurons? We show responses to similar stimuli for a variety of conventional model neurons in Fig. 1(b). None of the models reproduce the firing rate sensitivity of some *in vitro* neurons to input fluctuations above rheobase, the lowest DC current for which the neuron fires repetitively. For example, even if slow adaptation currents are added to the leaky-integrate-and-fire neuron, at high enough mean inputs the interspike interval equals the refractory period, yielding a firing rate that is insensitive to input fluctuations. What, then, is necessary for single-compartment model neurons to show noise sensitivity above rheobase?

To provide a framework within which to place these and previous results, we categorize neurons according to their *f–I* curves into three classes: Types A, B+, and B−. These types are distinct from those of Hodgkin's three classes (Hodgkin 1948; Rinzel and Ermentrout 1998; Gerstner and Kistler 2002; Izhikevich 2007), which focus on the bifurcations of neurons in response to noiseless currents. The *f–I* curves of Type A neurons show little sensitivity to input fluctuations above rheobase, as is the case for many standard single-compartment model neurons [Fig. 1(b)]. On the other hand, *f–I* curves from Type B neurons display sensitivity to input fluctuations throughout the neuron's dynamic range. We parse Type B into two categories depending on whether or

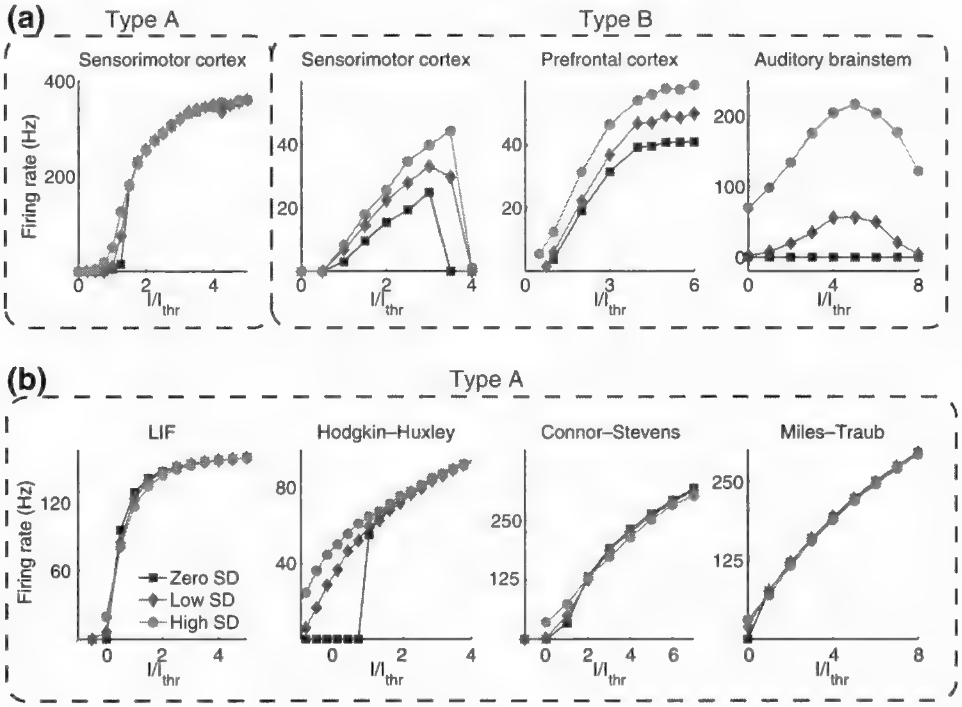


Fig. 1 The firing rates of *in vitro* neurons (*top row*) can display greater sensitivity to noise than those of conventional model neurons (*bottom row*). Type A neurons show little sensitivity to input fluctuations for large input means, while Type B neurons are sensitive throughout their physiological range. (**a**) Mean firing rate is plotted as a function of mean input current, where different traces represent different input SD with circles as highest input SD. From left to right, SD=[0, 0.7, 1.3], [0, 0.5, 1], [0.25, 0.75, 1.5], and [0, 1, 2] in units of I_{thr}, which was estimated to be the rheobase, or minimum zero-noise current needed to elicit spiking. From left to right I_{thr}=300, 400, 200, and 200 pA. From left to right, the first three neurons are from the neocortex (fast-spiking interneuron and two regular spiking pyramidal neurons) and the fourth is from the auditory brainstem (nucleus laminaris). Data from the sensorimotor cortex and auditory brainstem are courtesy of Matthew Higgs (Higgs *et al.* 2006), and data from the prefrontal cortex are courtesy of Michele Giugliano (Arsiero *et al.* 2007). (**b**) The mean firing rate of many single-compartment model neurons is insensitive to input SD for high input currents (i.e. above rheobase). From left to right, SD=[0, 0.5, 1], [0, 0.5, 1], [0, 1.5, 3], and [0, 1, 2] in units of I_{thr} with I_{thr}=40, 65, 100, and 25 nA/mm^2 respectively. The leaky integrate-and-fire neuron is a standard one-dimensional model often used for network modeling (Dayan and Abbott 2001), the classic Hodgkin–Huxley neuron models the squid giant axon (Hodgkin and Huxley 1952), the Connor–Stevens model is a prototypical Class I model based on a gastropod neuron (Connor and Stevens 1971), and the Miles–Traub model is based on hippocampal neurons (Ermentrout 1998)

not they fire repetitively to noiseless input. Type B− models do not fire repetitively unless noise is present in the input, while Type B+ models can. The transition from Type A to B- behavior has been explored previously (Lundstrom et al. 2008). We explore the causes leading to Type B+ behavior. We show how a separation of time scales between activation and recovery from spiking, such as might be induced by a voltage-dependent time constant for the recovery variable, promotes sensitivity to input fluctuations in low dimensional neuron models.

2 Methods

The two-dimensional model of Fig. 2 is described by the following equations, which are derived from the standard four-dimensional Hodgkin–Huxley (HH) model (Hodgkin and Huxley 1952; Koch 1999; Dayan and Abbott 2001; Gerstner and Kistler 2002) by eliminating the time dependence of m and letting h linearly depend on n, as in previous work (Rush and Rinzel 1995; Gerstner and Kistler 2002; Izhikevich 2007); we then altered the kinetics and conductances:

$$\begin{array}{l} C\frac{dV}{dt} = -G_{\mathrm{Na}} m_\infty^3 h(V - E_{\mathrm{Na}}) - G_{\mathrm{K}} n^4 (V - E_{\mathrm{K}}) - G_{\mathrm{Leak}}(V - E_{\mathrm{Leak}}) + I \\ \tau\frac{dn}{dt} = n_\infty - n \end{array} \qquad (1)$$

where

$$\begin{array}{l} m_\infty = 1/(1 + \exp((-40 - V)/k_m)) \\ h = 0.89 - 1.1n \\ n_\infty = 1/(1 + \exp((V_n - V)/k_n)) \end{array}, \qquad (2)$$

with parameters C=1 nF/mm^2, G_{Na}=50 mS/cm^2, G_{K}= 36 mS/cm^2, G_{Leak}=5 mS/cm^2, E_{Na}=50 mV, E_{K}=−77 mV, E_{Leak}=−54 mV, k_m=7, V_n=−45, k_n=15, and τ=5 or 100 ms (except as noted). This 2D model loses stability via a subcritical Hopf bifurcation, and can be termed a Class II model (Rinzel and Ermentrout 1998; Izhikevich 2007). With modifed parameters, G_{Leak}=15 ms/cm^2, V_n=−30 mV, and k_n=5, the model is Class I and loses stability via a saddle-node on invariant circle bifurcation. Bifurcation analyses were done using XPPAUT (developed by Bard Ermentrout). When τ was not constant, its voltage dependence was taken to be (Izhikevich 2007):

$$\tau = C_{\mathrm{Base}} + C_{\mathrm{Amp}} \exp\left(-(V_{\max} - V)^2 / \Delta^2\right), \qquad (3)$$

with the parameters C_{Base}, C_{Amp}, $V_{\max}$, and Δ. For models of a similar form, see Izhikevich (2007). To implement slow sodium channel inactivation, an extra gating variable s was added to the sodium current so that it became:

$$I_{\mathrm{Na}} = G_{\mathrm{Na}} m^3 hs(V - E_{\mathrm{Na}}) \qquad (4)$$

This sodium current was incorporated into the previous Class I 2D HH model. The dynamics of the variables n and h remained unchanged except that τ=3 ms. The slower gate s was identical to h except that its time constant was governed by Eq. (3) with parameters C_{Base}=50 ms, C_{Amp}= 2,000 ms, $V_{\max}$=−40 mV and Δ=5 mV.

Importantly for this study, input fluctuations were fast compared to the dynamics of spike recovery and time between spikes. The external input current I to the neuron was 1-ms exponentially filtered Gaussian noise (i.e. equivalent to an Ornstein–Uhlenbeck process) with the means and standard deviations as specified. Results were qualitatively identical with a 5-ms filter. Model equations were integrated using a fourth-order Runge–Kutta solver with a 0.02 ms fixed time step. Spikes were counted as upward voltage crossings at −20 mV if the mean of the previous 1 ms of voltages was less than −40 mV.

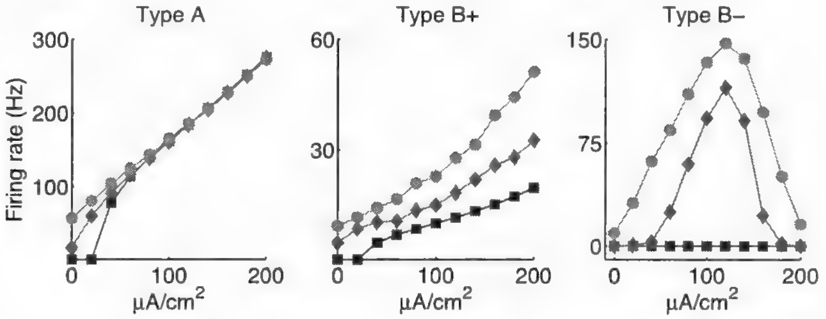


Fig. 2 A two-dimensional modified and reduced Hodgkin-Huxley (HH) model neuron can show all three types of behavior. Type A is similar to the standard HH model and is insensitive to input SD for high currents. In contrast, Type B+ is sensitive to input SD throughout the dynamic range and fires repetitively to inputs with SD=0. Type B− models never fire repetitively when input SD=0 and never undergo a bifurcation from stable fixed point to limit cycle. For the three models, G_{Na} and τ were [50, 50, 15] mS/cm^2 and [5, 100, 5] ms, respectively. Input SD was [0, 10, 20] µA/cm^2. Other parameters were as given in the Section 2

3 Results

Many single-compartmental biophysical model neurons, implemented with their standard parameters, display *f–I* curves best described as Type A [Fig. 1(b)]: their firing rates are not sensitive to input fluctuations above rheobase. Here, we are interested in model neurons that display sensitivity to noise above rheobase until depolarization block, which we have called Type B. We subclassify Type B neurons as Type B− when they do not fire repetitively to noiseless current regardless of its magnitude, and Type B+ if they do for some magnitude. In previous work (Lundstrom et al. 2008), we demonstrated that the alteration of a model's conductance ratio, e.g. G_{Na}/G_K, can lead to Type B− *f–I* curves. This behavior is generated by neuronal dynamical systems with fixed points that remain stable regardless of input mean. Thus, these neurons never fire repetitively at steady state in response to noiseless input. We focus here on Type B+ neurons whose firing rates are sensitive to input fluctuations throughout the dynamic range and which can fire repetitively at steady state to noiseless input.

3.1 2D model demonstrating three types of *f–I* curves

We begin with a two-dimensional model, similar to the Hodgkin–Huxley neuron, that can demonstrate three types of *f–I* curves: Type A, B+, and B− (Fig. 2). We wish to identify the specific characteristics of the differential equations describing the neuronal dynamics that lead to the generation of Type A vs. B+ behavior. For 2D dynamical systems, these characteristics can be explored geometrically using phase portraits. To do this, we reduced the standard 4D HH model to two dimensions by eliminating the time dependence of *m* and letting *h* linearly depend on *n* (Izhikevich 2007); we slightly altered the kinetics and conductances. We then examined 2D model trajectories in the phase plane for each of the neuron types.

Two-dimensional dynamical systems can be analyzed by examining a phase portrait, which is a plot of one dynamical variable against the other (Strogatz 1994; Gerstner and Kistler 2002; Izhikevich 2007). In this case, the model has a fast positive feedback variable *V* and a slow negative feedback variable *n*. *V* is the model's membrane voltage, while *n* is a combined variable, representing sodium channel inactivation as well as potassium activation. As the membrane voltage *V* spikes in time, the neuron's trajectory travels counter-clockwise around the phase plane (Fig. 3). The upswing and downswing of the action potential (dashed lines) correspond to the left-to-right and right-to-left trajectory jumps, respectively, between the arms of the *V*-nullcline (thick solid line). The *V*-nullcline and *n*-nullcline (fine solid line) correspond to points on the phase plane where $dV/dt=0$ and $dn/dt=0$, respectively.

3.2 The effect of increasing time scale separation

From the behavior of the trajectories of this two-variable (*V* and *n*) model in the phase plane, we can relate the

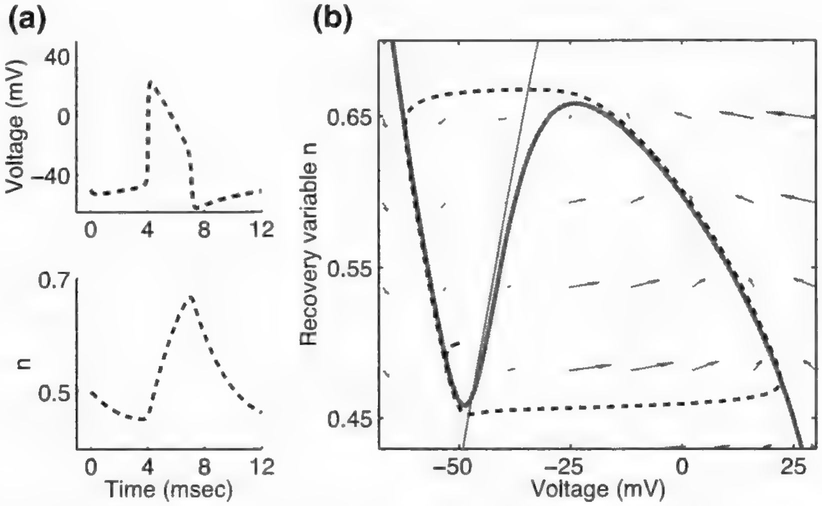


Fig. 3 The action potential in the phase portrait. (**a**) The membrane voltage and recovery variable *n* change as a function of time and (**b**) can be plotted against each other in a phase portrait. The membrane voltage as a function of time shows a sharp upswing that corresponds to a rightward, horizontal movement in the 2D phase plane of voltage, the activation variable *V*, and the recovery variable *n*. Dashed lines represent the neuron's trajectory, the N-shaped solid line is the *V*-nullcline, the solid straight line is the *n*-nullcline, and arrows represent the *flow field* with vectors given by $(dV/dt, dn/dt)$ at each point (V, n). Data are from the Type A model in Fig. 2 with stimulus mean $I = 50\ \mu A/cm^2$

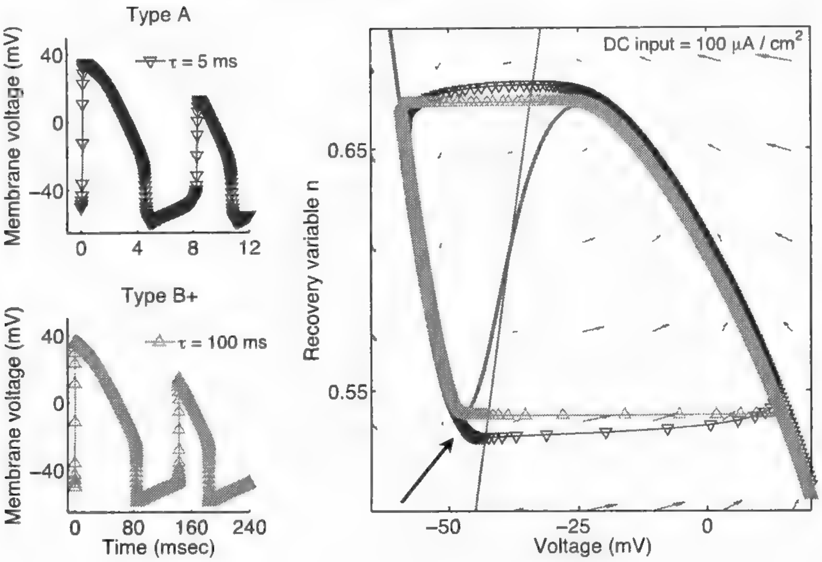


Fig. 4 As the recovery time scale τ increases, the upswing of the action potential (large, black arrow) occurs closer to the local minimum of the *V*-nullcline (curved green line). Changing the recovery time scale τ does not alter the nullclines but rather affects how neuronal trajectories move in the phase plane, which can also be seen in voltage vs. time plots (left; notice the different x-axis scale). The direction and magnitude of the small, green arrows represent the relative magnitudes of *dV/dt* and *dn/dt* for the *(V, n)* points in the phase portrait. Notice that trajectories overshoot the local minimum of the left arm of the *V*-nullcline (large arrow) only when τ is small

computational behavior of the models to properties of the dynamical system. We considered two models that differed only in the value of the recovery variable time constant τ: a short time scale model with τ=5 ms, and a model with τ= 100 ms (see Methods). Figure 4 shows responses of both models to an input with a standard deviation of zero and an input mean of μ=100 μA/cm^2. Without input fluctuations, the neuron spikes regularly with a stereotyped voltage trajectory. In other words, during the spike upswing the value of n is the same from spike to spike.

The time constant τ controls the speed of the slow variable n in this model, while the speed of the fast variable V is governed by G_{Na}, G_{K}, and G_{Leak}, biophysical conductance parameters describing the density of sodium, potassium, and leak channels in the membrane. Because the ionic conductances are constant, τ represents the time scale separation in the model: as τ increases, the time scale separation between the two variables increases. For increasing τ, the spikes have a similar shape but are wider. Spikes of this width are clearly not physiological; this is a limitation of an increased time scale separation in 2D systems that we address later. The phase plane representation shows that as τ decreases, trajectories start to overshoot the left and right bends of the V-nullcline (Fig. 4).

With a noisy input, voltage trajectories are less stereotyped and the intervals between spikes are irregular (Fig. 5). In the model with small τ, noise does not change the mean value of n during spike upswings and downswings but only adds variability [Fig. 5(a)]. Thus, when there is little separation of time scales, input fluctuations do not alter the mean firing rate, as seen in Fig. 2, but only spike timing. However, in the model with large τ, input noise shortens the average oscillation period, increasing the mean firing rate [Fig. 5(b)]. Although noise does not appreciably change $\langle n \rangle$ on the spike downswing, it does on average increase n during the upswing. When τ is large, trajectories travel more slowly down the left arm of the V-nullcline. This allows input fluctuations the opportunity to initiate a spike sooner by pushing the neuron past its spiking threshold, which is approximately given by the upward-turning middle arm of the V-nullcline (Hong et al. 2007; Izhikevich 2007). Although this threshold is two-dimensional, once the neuron begins to fire, n stays relatively constant during the fast upswing of the action potential.

3.3 Defining an instantaneous potential barrier

Since noise causes the trajectory to jump across the threshold sooner, this implies that there is a barrier that prevents crossing in the absence of noise. To gain insight into how noise drives spiking, we examined how noise-driven trajectories escape over a barrier. Consider a simple 1D model as in Fig. 6(a), where input fluctuations of typical scale σ cause trajectories to move in the voltage V dimension, such that sometimes the trajectory can overcome an instantaneous potential barrier ΔU located at a threshold for spiking. This picture is reminiscent of problems in physics and chemistry wherein the activation rate is determined by the size of an energy barrier and the temperature and is given by the Arrhenius rate or Kramer's

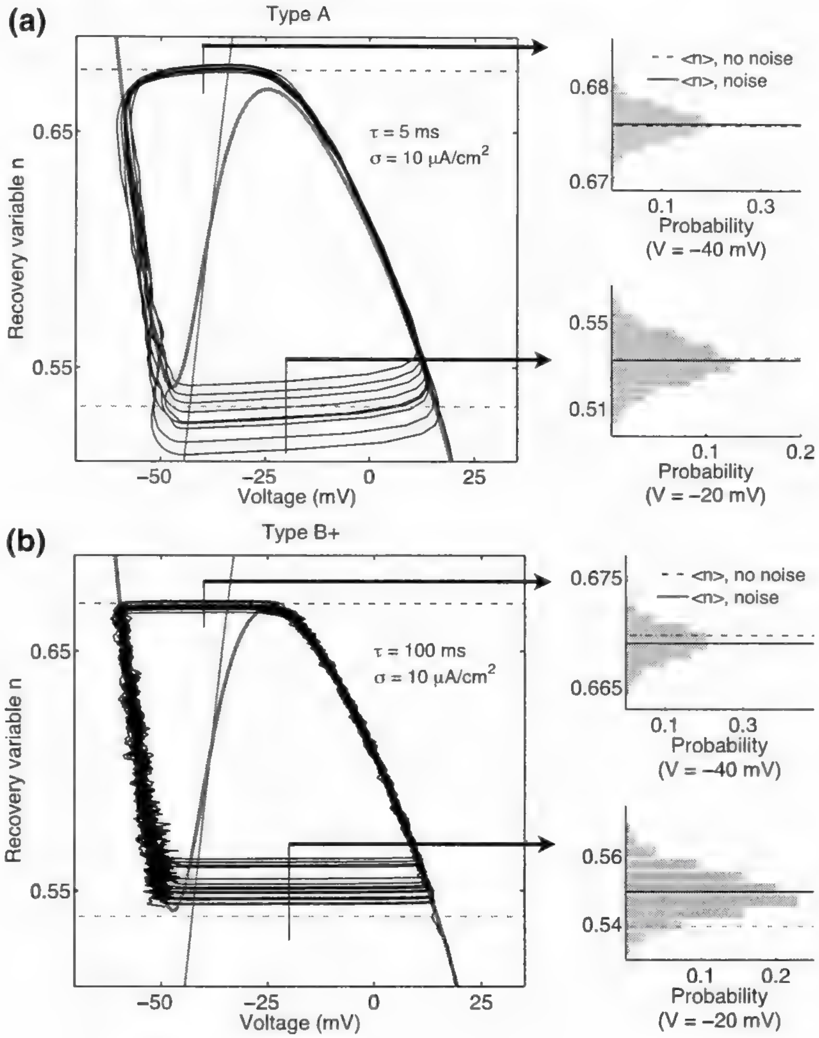


Fig. 5 Input fluctuations do not change the mean firing rate when τ is small, but increase firing rate when τ is large. (**a**) When τ is small (5 ms), input fluctuations (SD=10 μA/cm^2) increase the variance of n during the up- and downswings of the action potential, but do not alter the mean value of n. Histograms are shown at right during action potential upswing (V=-20 mV) and downswing (V=-40 mV), as indicated by the vertical black lines on the phase portraits. The dashed lines represent the value of n when input SD=0 mV, while the solid lines show the mean values of the data. The dashed and solid lines are nearly the same; the neuron's firing rate does not change with increased input SD, but spiking becomes irregular. (**b**) When τ is large (100 ms), increasing input SD leads to an increasing mean firing rate. Although the mean value of n during the action potential downswing does not appreciably change, during the upswing $<n>$ increases, since on average the input SD causes the neuron to spike sooner, i.e. before n has returned to the minimum. The input current I had a mean of 100 μA/cm^2

escape rate equation, $r \sim \exp(-\Delta U/kT)$. In our case, thermal energy kT is replaced by a factor proportional to the variance of the driving current fluctuations, σ^2.

DeVille et al. (2005) developed this framework for a Fitzhugh–Nagumo excitable system in the limit when noise σ is very small and τ is very large. We wish to apply this approach to situations where the values of input noise σ and the separation of time scales τ are relevant for physiological neurons. First, we test this relationship for less extreme values of σ and τ.

To find ΔU, we need the relevant potential. In general, the *potential* is defined as a scalar function with the property that its negative derivative is the force (Arfken and Weber 1995). When the separation of time scales is large, we can consider n to be fixed in the equation of motion for V, Eq. (5), so that the relevant force is $-dV/dt$ and the problem reduces to a series of one-dimensional cases for various n. We therefore define the negative integral of dV/dt as an instantaneous potential landscape, as in Fig. 6(a), for each value of n. Specifically, we consider:

$$\frac{dV}{dt} = \left(\begin{array}{r} -G_{\mathrm{Na}} m_{\infty}^{3}(V) h(n)(V - E_{\mathrm{Na}}) - G_{\mathrm{K}} n^{4}(V - E_{\mathrm{K}}) \\ - G_{\mathrm{Leak}}(V - E_{\mathrm{Leak}}) + I \end{array} \right) / C, \tag{5}$$

which can be rewritten as:

$$k_{\mathrm{c}} = f(V, n), \tag{6}$$

where the k_{c}=0 contour is the nullcline. As in DeVille et al. (2005), we seek a function U such that

$$\frac{\partial U(V, n)}{\partial V} = -f(V, n). \tag{7}$$

Integrating Eq. (6) with respect to V, we obtain $U(V)$ for a given value of n. This produces a double-well potential

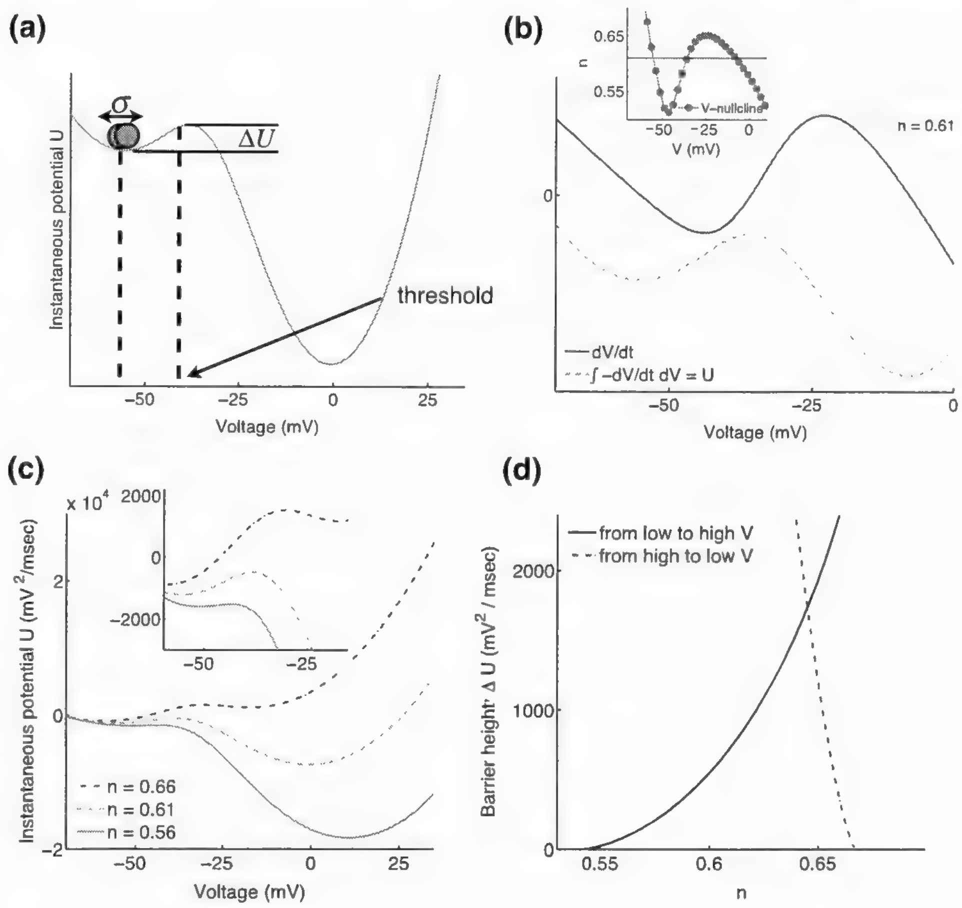


Fig. 6 Input fluctuations can shorten interspike intervals by causing the neuron's trajectory to cross an instantaneous potential barrier. (**a**) A simple 1D model relates spike initiation to crossing an energy barrier. (**b**) The instantaneous potential landscape can be found by integrating $-dV/dt$ (solid, blue line) with respect to V for constant n, here represented by the solid black line in the inset. The result of the integral is represented by the dashed, green line. (**c**) The instantaneous potential landscape changes as n changes, and potentials are shown for three specific values of n, which represent three different slices through the inset of (b). The action potential upswing and downswing occur at approximately $n=0.56$ and $n=0.66$, respectively. The middle hump is the barrier related to the spiking threshold. Units are defined up to a constant. (**d**) The barrier height from low to high V is the potential U at the middle hump subtracted from the potential U for the left hump. As n increases, the barrier height for the rapid depolarization of the spike upswing increases

for each n, where the local minima and maxima of $U(V)$ occur where the V-nullcline passes through that value of n. For an N-shaped nullcline, this leads to two minima and one maximum [Fig. 6(b)]. The shape of this instantaneous potential changes with n. During the spike upswing, the trajectory "rolls" down the potential to the right [Fig. 6(c, d), solid lines]. Then, as membrane voltages remain high, n increases until eventually the trajectory "rolls" down the potential to the left (Fig. 6(c, d), dash-dotted lines). When the separation of time scales is large, the potential landscape is quasistatic, and spikes can occur sooner due to noise-driven jumps over the potential barrier.

3.4 Relating τ and ΔU

Given this expression for U, for every n there exists a barrier height ΔU. By the Arrhenius, or Kramer's escape rate, equation, the time scale $\lambda(n)$, or inverse rate, for jumping the barrier for a given input σ is given by:

$$\lambda(n) = k\exp\left(\Delta U(n)/m\sigma^2\right), \tag{8}$$

where m is a constant. As in DeVille et al. (2005), for low noise levels the trajectory following a spike moves down the left arm of the V-nullcline with a time scale determined by the recovery time constant τ. While following this trajectory, the jumping rate $\lambda^{-1}(n)$ changes with n. The system becomes susceptible to barrier crossing when $\lambda(n)$ is of same order as the recovery time scale τ. Therefore, we expect that average barrier crossing behavior of the system is characterized by

$$\tau = \left\langle k\exp\left(\Delta U(n)/m\sigma^2\right)\right\rangle, \tag{9}$$

where the brackets indicate an average over many spikes. Furthermore, since the variance of n at barrier crossing is small, and the value of n is nearly constant during each spike upswing and downswing, as seen in Fig. 5, we can simplify Eq. (9), without significant loss of accuracy, to:

$$\tau \cong k\exp\left(\Delta U(\langle n\rangle)/m\sigma^2\right). \tag{10}$$

$\Delta U(\langle n\rangle)$ is determined by evaluating the potential difference at $\langle n\rangle$ for either the spike upswing (fast depolarization) or spike downswing (fast repolarization). If Eq. (10) accurately describes this system, we should find that $\Delta U(\langle n\rangle)$ is linearly related to $\ln(\tau)$. Using the values of $\langle n\rangle$ that correspond to a given τ [Fig. 7(a)], we find that ΔU and τ are indeed exponentially related for a given σ

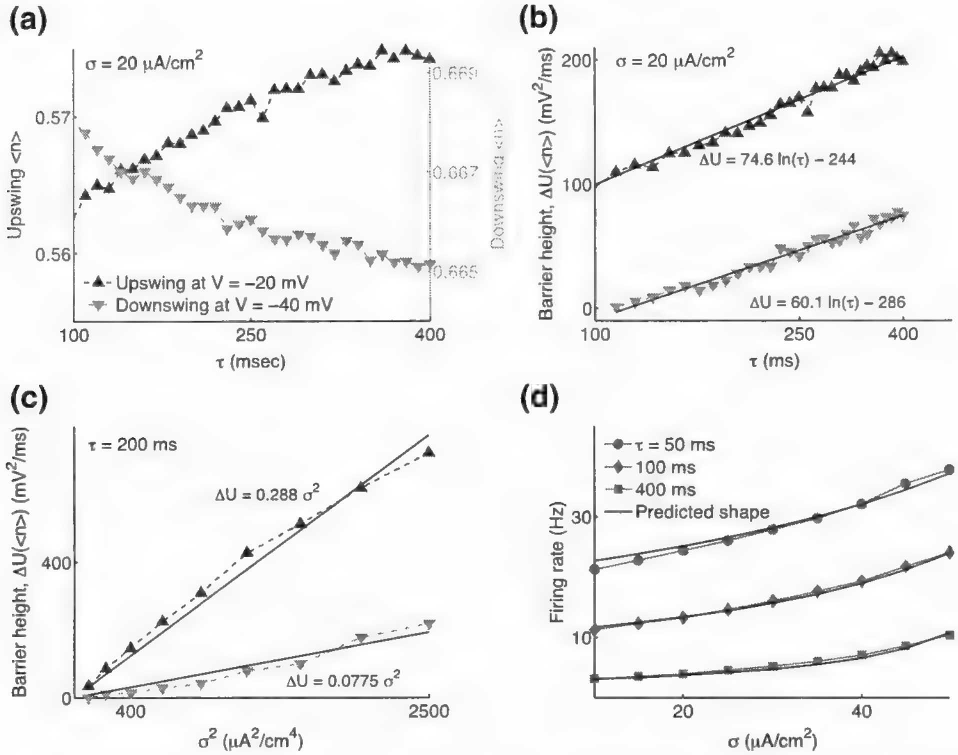


Fig. 7 The ability of input fluctuations to cause an increase in firing rate is fit by an exponential relationship with the time scale separation. (**a**) As recovery time constant increases, action potentials are initiated and terminated sooner, i.e. upswings are at higher mean values $\langle n \rangle$ and downswings are at lower $\langle n \rangle$. For each n there exists a barrier height ΔU, as illustrated in Fig. 6. The upswing $\langle n \rangle$ and downswing $\langle n \rangle$ were measured at V=−20 and V=−40 mV, respectively, as in Fig. 5, and the symbols remain the same for the next two panels. (**b**) The mean barrier height $\langle \Delta U \rangle$ is exponentially related to the separation of time scales, (**c**) and linearly related to input variance σ^2. (d) Firing rate increases with variance, as seen from simulation data (*circles*, *diamonds*, and *squares*). These firing rates can also be calculated as a function of σ with the parameter τ=[50, 100, 400] ms using Eq. (15). The input current I had a mean of 100 μA/cm^2

[Fig. 7(b)], as suggested by the simple 1D picture of Fig. 6(a). The results in Fig. 7(b) and (c) can be summarized by the following equation:

$$\Delta U(\sigma, \tau) = (A \ln(\tau) + B)\sigma^2, \tag{11}$$

with A and B constant, and where ΔU and constants differ for the upswing and downswing.

Thus, the simple picture of Fig. 6(a) holds approximately even away from the limits of very large time scale separation and low noise. We now wish to predict the mean firing rate as a function of σ and τ. Equation (11) relates σ and τ with ΔU. We next relate the oscillation period T to n_{up}, the value of n when membrane voltage rapidly depolarizes during the spike upswing. Finally, we relate n_{up} to ΔU.

We began by examining n and T. As seen in Fig. 4, for large τ the trajectory spends essentially all of its time slowly moving along the left and right arms of the V-nullcline. Given this, one can find an approximate analytical expression for $T(n)$ directly from the model equations. Along the V-nullcline, dV/dt=0, and thus motion along the nullcline, i.e. along $f(V, n)$=0, can be described by dynamics due to dn/dt. In this case, the equation of motion is:

$$\tau \frac{dn}{dt} = n_\infty(V) - n.$$

The period of the trajectory, i.e. the interspike interval, is dominated by the time spent traveling on the left and right arms of the V-nullcline,

$$T = \tau \left(\int\limits_L dn \left(\frac{1}{n_\infty(V) - n} \right) + \int\limits_R dn \left(\frac{1}{n_\infty(V) - n} \right) \right), \tag{12}$$

where the bounds of integration are given by n_{up} and n_{down} for the left and right arms. One can numerically integrate Eq. (12) to find $T(n_{\text{up}}, n_{\text{down}})$.

We can, however, take a simpler approach based on the following observations. Due to the geometry of the nullclines, changes in n_{down} are small relative to those in

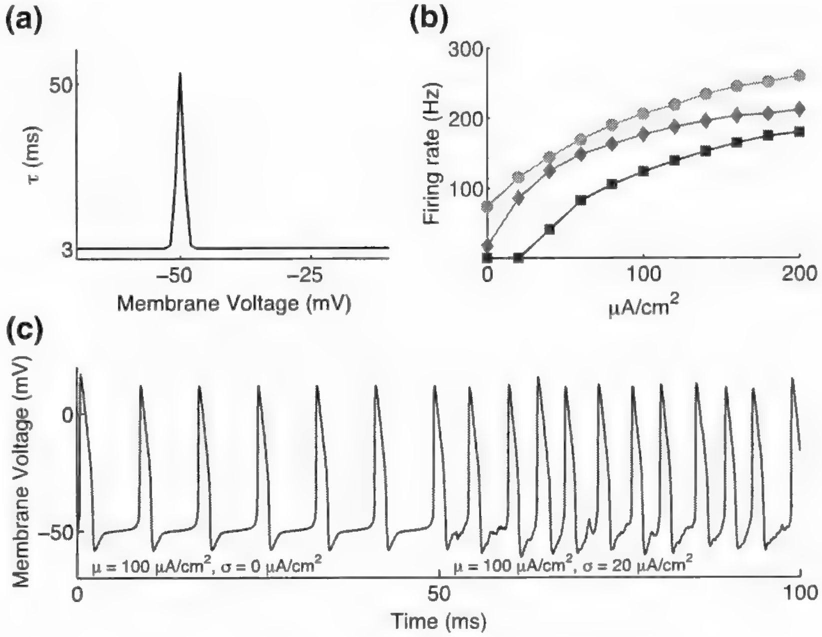


Fig. 8 Addition of voltage dependence to the slow time scale of a 2D Hodgkin–Huxley model gives Type B+ behavior. (**a**) The time constant of the recovery variable is taken to be large only at peri-threshold voltages. $C_{\mathrm{Base}}=3$ ms, $C_{\mathrm{Amp}}=50$ ms, $V_{\max}=-50$ mV, $\Delta=1$ mV (see Section 2). (**b**) This allows input fluctuations to increase average firing rate. SD=0, 10, 20 μA/cm² for squares, diamonds, and circles, respectively. (**c**) Membrane voltages corresponding to two points on the *f*–*I* curve as indicated. Spike widths are narrow

n_{up} (Fig. 5), and the firing rate is relatively insensitive to n_{down}. We therefore consider only the dependence on n_{up}. Further, focusing on the left branch, the *V*-nullcline and the *n*-nullcline $n_\infty(V)$ can be approximated as linear in this region, as suggested by the phase plots (e.g. Fig. 5). From the flow field, as depicted by the arrows in Figs. 3 and 4, it can be seen that the rate at which *n* changes with time decreases as the trajectory comes down the left branch. Thus, the time spent along the left branch is:

$$T_{\mathrm{L}} = \tau \int_{n_{\mathrm{down}}}^{n_{\mathrm{up}}} \frac{1}{k_1 - k_2 n} dn = \tau \ln\left(\frac{k_1 - k_2 n_{\mathrm{up}}}{k_1 - k_2 n_{\mathrm{down}}}\right)$$
$$= \tau \ln\left(1 + \frac{\left(n_{\mathrm{down}} - n_{\mathrm{up}}\right)}{k_1/k_2 - n_{\mathrm{down}}}\right),$$

where k_1 and k_2 are positive constants and $k_1 >> k_2$. A similar expression gives the time along the right branch. Using $\ln(1+x) \approx x$ for small *x*, the oscillation time is approximated by

$$T \approx \tau\left(-\alpha n_{\mathrm{up}} + \beta\right), \tag{13}$$

where $n_{\mathrm{down}} - n_{\mathrm{up}}$ is small relative to n_{down} and α and β are positive constants.

Now, we relate ΔU and *n*. For the spiking barrier, i.e. for ΔU_{up}, this is accomplished by finding $V_1^*(n)$ and $V_2^*(n)$, the voltages for which $dV/dt=0$ on the left and middle arms given *n*, and substituting them into $U(V)$, which is defined by Eq. (7). This gives $\Delta U_{\mathrm{up}} = U\left(V_2^*\right) - U\left(V_1^*\right)$ as in Fig. 6(d). In general, $\Delta U_{\mathrm{up}}(n)$ and $\Delta U_{\mathrm{down}}(n)$ are complicated functions that depend on the widths between the left and middle branches and right and middle branches of the *V*-nullcline, respectively. However, as can be seen from the shape of the *V*-nullcline, they are monotonic, and numerically we find $\sqrt{\Delta U(n)} \propto n$ to be a good approximation. We now have the necessary equations to calculate firing rate:

$$\begin{aligned} &\Delta U = (A_1 \ln \tau - B_1)\sigma^2 \\ &\sqrt{\Delta U} = A_2 n - B_2 \\ &T = (-A_3 n + B_3)\tau \end{aligned} \quad , \tag{14}$$

where A_i and B_i are positive constants. Straightforward algebra yields an equation for the mean period, *T*:

$$T = \tau \frac{-A_3}{A_2}\left(\sqrt{(A_1 \ln \tau - B_1)\sigma^2} + B_2\right) + \tau B_3, \tag{15}$$

and thus $T \propto (\text{const}-\sigma)$ for a given τ. We find very good agreement between Eq. (15) and the simulated data [Fig. 7(d)].

3.5 Type B+ behavior with narrow spike widths

In this 2D model, large τ leads to very wide spikes (e.g. Fig. 4), so this regime of a separation of time scales may seem physiologically irrelevant. To achieve B+ behavior with narrow spike widths in this model, one can manipulate the geometry of the nullclines around the left knee of the *V*-nullcline (Fig. 4, black arrow). For example, the model of Eq. (2) with k_m=5 and k_n=23 both has a sharper left knee of the *V*-nullcline and an *n*-nullcline that is reduced in slope, slowing spike trajectories at the left knee. This model exhibits Type B+ behavior with τ=5 ms and spike widths of approximately 5 ms. Increasing the dimensionality of the model is an alternative approach, allowing separation of the slow time constant from the spike recovery process. For example, one can add long time-scale adaptation through a

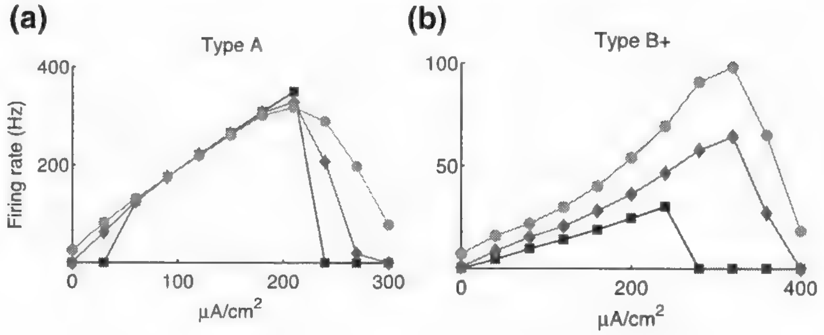


Fig. 9 A Class I 2D Hodgkin–Huxley model can show either Type A or Type B+ behavior. The previous Class II model is now Class I (see Section 2) with (**a**) τ=5 ms, or (**b**) τ=100 ms. SD=0, 10, 20 μA/cm^2 for squares, diamonds, and circles, respectively

slow spike-dependent current with a constant time constant, thus altering the shape of the V-nullcline as input current is increased.

Until now, we have considered the slow time constant of the recovery variable to be constant. However, a straightforward way to achieve narrow spikes in a 2D model is by allowing this time scale to be voltage dependent. When the recovery time constant becomes very large for peri-threshold voltages [Fig. 8(a)], spike trajectories slow down during depolarization, allowing the influence of noise. Thus, average firing rates become sensitive to input fluctuations [Fig. 8(b)], but spike widths are nonetheless narrow [Fig. 8(c)]. The details of the voltage dependence can strongly influence the details of shapes of the f–I curves.

3.6 Relation to Hodgkin's Classification

Our classification of neurons into types A and B is distinct from Hodgkin's Class I/II/III categories (Hodgkin 1948). Hodgkin classified neuronal response to noiseless injected current according to whether action potentials could be generated with an arbitrarily low frequency (Class I), with a certain minimum frequency with firing rates that were rather insensitive to applied current (Class II), or with only very strong current or not at all (Class III; Hodgkin 1948; Gerstner and Kistler 2002; Izhikevich 2007). This classification scheme focuses on how neurons respond to inputs in the absence of input fluctuations, and has been identified with the bifurcation structure of the neuronal dynamical system (Rinzel and Ermentrout 1998; Izhikevich 2007), while our approach emphasizes neuronal sensitivity to input fluctuations.

Although Type B− and Class III neurons are equivalent for excitable neurons (Hong et al. 2008; Lundstrom et al. 2008), there is no apparent correspondence between Classes I and II and Types A and B+. The 2D reduced HH models (Types A and B+) are Class II models, as verified by bifurcation analyses showing a loss of stability via a subcritical Hopf bifurcation. With the choice of parameters G_{Leak}=15 ms/cm^2,

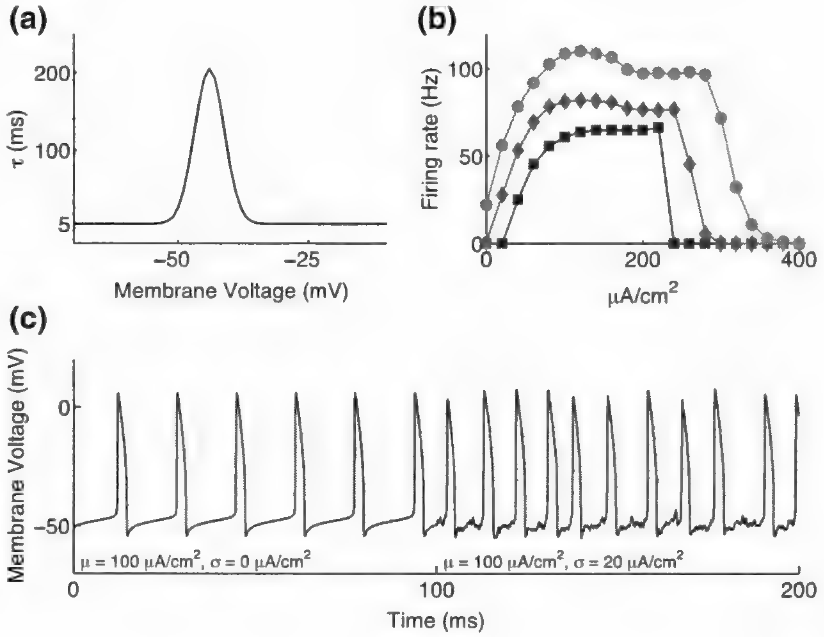


Fig. 10 Addition of a voltage-dependent slow time scale to a Class I 2D HH neuron model. (**a**) The time constant of the recovery variable is large only at peri-threshold voltages. C_{Base}= 5 ms, C_{Amp}=200 ms, V_{max}= −44 mV, Δ=4 mV (see Section 2). (**b**) SD=0, 10, 20 μA/cm^2 for *squares*, *diamonds*, and *circles*, respectively. (**c**) Membrane voltages corresponding to two points on the f–I curve as indicated

V_n=−30 mV, and k_n=5, the reduced models become Class I, Fig. 9. Their corresponding *f–I* curves are qualitatively identical to those of the Class II models, suggesting that the particular way in which the model loses stability does not dramatically influence its sensitivity to input fluctuations over its dynamic range. Similar results were obtained using the Morris-Lecar model (Morris and Lecar 1981) implemented with Class I dynamics (Rinzel and Ermentrout 1998; Tateno and Pakdaman 2004). This is in agreement with previous work showing that Class I and Class II Morris–Lecar models respond similarly to noisy inputs (Tateno and Pakdaman 2004), although other work has highlighted the models' different transient responses to fluctuating inputs near firing threshold (Gutkin and Ermentrout 1998; Robinson and Harsch 2002). As with the Class II models, adding voltage dependence to the recovery time constant of Class I neurons can give rise to Type B+ behavior. In Fig. 10, the time constant is larger for peri-threshold values, and the shape of the *f–I* curves depends on the particular form of voltage dependence, especially for inputs with large input fluctuations.

3.7 Higher dimensional models and adaptation

For these 2D models, a voltage-dependent recovery variable that slows greatly near threshold allows the neuron to be sensitive to input fluctuations while maintaining narrow spike widths. While in 2D models the time scales for voltage reset and for recovery from spiking are combined into a single variable, higher-dimensional models may have more than two time scales. Thus, slow recovery from sodium inactivation (Fleidervish et al. 1996) may serve a similar function. In Fig. 11, we added a sodium inactivation gating variable to the Class I 2D HH model (see Section 2). This 3D model has three time scales: a fast time scale for membrane voltage, a slow time scale for the linearly related recovery variables *n* and *h*, and a yet slower, voltage-dependent time scale for the slow sodium inactivation variable *s*. This model also displays Type B+ behavior, over a wider parameter range than the 2D model of Fig. 10. The repolarization of spikes in Fig. 11(c) is governed by the intermediate time scale of *n* and *h*, while the slow depolarization just prior to spiking threshold is governed by the slowest time scale of *s*. When *s* has a fixed time constant, the effect of slow sodium inactivation in this model is to modify the shape of the *V*-nullcline differentially as firing rate increases; as for 2D models, this can also facilitate B+ behavior but the effect is not strong.

The important required effect is that the trajectories remain just below threshold for an extended time, allowing input fluctuations to initiate spiking sooner on average. For phase portraits such as Fig. 5, this corresponds to trajectories spending an extended time near the left knee of the *V*-nullcline. Lower energy barriers ΔU occur when the system is held near threshold, and when there is a smaller distance between the left and center arms of the *V*-nullcline, Fig. 6. Thus, slow sodium inactivation, in addition to causing spike frequency adaptation (Fleidervish et al. 1996), may facilitate sensitivity to input fluctuations as in Type B+ behavior.

4 Discussion

We set out to understand, using the simplest models possible, what is required to produce the experimentally observed

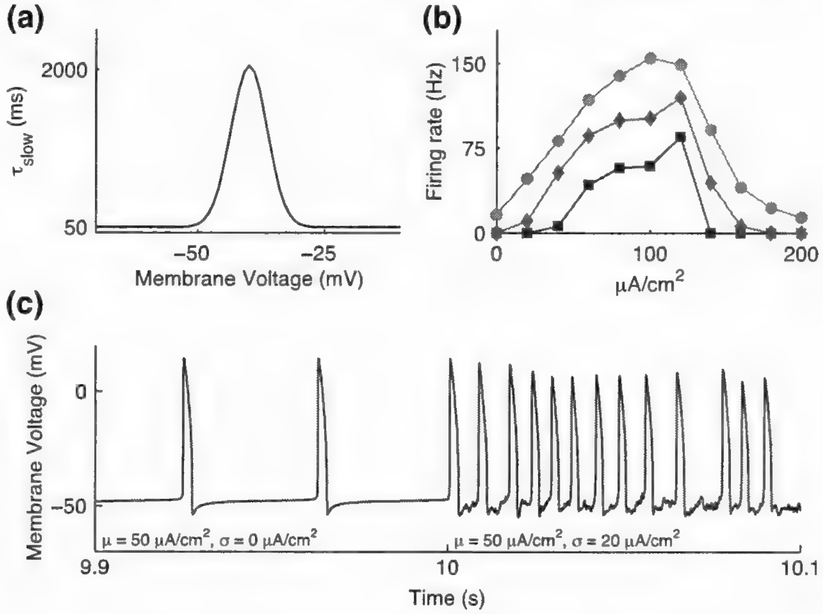


Fig. 11 Addition of slow sodium inactivation to 2D HH neuron model. (**a**) The time constant of the slow sodium inactivation gate is large only at peri-threshold voltages. C_{Base}=50 ms, C_{Amp}=2,000 ms, V_{max}=−40 mV, Δ=5 mV (see Section 2). (**b**) SD=0, 10, 20 μA/cm² for *squares*, *diamonds*, and *circles*, respectively. (**c**) Membrane voltages corresponding to two points on the *f–I* curve as indicated. Slow depolarization near threshold voltages is controlled by slow sodium inactivation

diversity of *f–I* curves in response to input mean and variance (Chance et al. 2002; Fellous et al. 2003; Rauch et al. 2003; Higgs et al. 2006; Arsiero et al. 2007), as in Fig. 1(a). Many single-compartment neuron models display converging *f–I* curves and low sensitivity to input fluctuations at high input mean currents [Fig. 1(b)]; we call these Type A neurons. However, there are clear experimental examples of neurons whose *f–I* curves show sensitivity to input fluctuations throughout their dynamic range [Fig. 1(a)], which we term Type B. Neuronal sensitivity to input fluctuations may provide an additional communication channel for neural coding (Fairhall et al. 2001), implement a type of noise-driven gain control (Chance et al. 2002; Higgs et al. 2006), underlie the ability of neurons to perform coincidence detection (Konig et al. 1996; Slee et al. 2005; Higgs et al. 2006), and promote persistent activity in networks (Arsiero et al. 2007). These computational regimes are distinct from Hodgkin's classification scheme; for example, a Type B neuron may be either Class I or II.

We parse Type B into two subtypes: B− and B+. B-neurons do not repetitively fire at steady state to noiseless current, regardless of how large, and are similar to differentiating neurons that respond primarily to input changes as seen in the auditory brainstem (Slee et al. 2005; Higgs et al. 2006). Previous work (Lundstrom et al. 2008) has explored the underlying causes of B− responses: these neurons always have a stable fixed point, meaning that input variance is an essential cause of spiking. Further, there is a well-defined boundary in conductance space between Type A and B−, such that upon lowering the conductance G_{Na}, a neuron can suddenly switch from Type A to B−. This boundary is planar in the 3D space of (G_{Na}, G_{K}, G_{Leak}), where an equation of the form $G_{\mathrm{Na}}-AG_{\mathrm{K}}-BG_{\mathrm{Leak}}=0$ describes the boundary. The coefficients A and B can be determined analytically from model equations. Thus, it is a ratio of depolarizing and hyperpolarizing conductances that determines when a neuron is Type A or B−.

On the other hand, B+ neurons can repetitively fire to noiseless current and are more similar to some cortical neurons (Rauch et al. 2003; Higgs et al. 2006; Arsiero et al. 2007). Here, we link Type B+ behavior to two factors: the shape of the *V*-nullcline, which determines the height of an effective barrier to spiking, and a separation of time scales between an activation and recovery variable. Implicitly, the time scale of input fluctuations, which may be influenced by synaptic filtering and membrane time constants, must also be fast relative to that of the recovery variable. Our classification is based on sensitivity to input fluctuations, which is somewhat arbitrary. Here, we have classified neurons as sensitive to fluctuations, or Type B, when their firing rates vary by more than a few percent as input fluctuations are increased from zero to physiologically realistic levels (i.e. such that noise-induced membrane voltage fluctuations are small compared to spike height). The transition from Type A to B for model neurons can be gradual, thus potentially creating difficulties with regard to classification. However, many *in vitro* and model neurons, as in Fig. 1, clearly fall into one of these types, emphasizing distinct differences in sensitivity to input fluctuations.

Although we initially focused on 2D systems, our results apply to 3D neuron models (Fig. 11), and thus may also be relevant for more complex experimentally recorded neurons. Slow sodium inactivation allows the voltage trajectory to stay near threshold for a prolonged amount of time, which gives rise to Type B+ behavior. This may provide a theoretical underpinning for *in vitro* data showing Type B+ behavior in response to somatically injected current Gaussian noise (Higgs et al. 2006; Arsiero et al. 2007). In these studies, slow adaptive mechanisms were implicated, both slow sodium inactivation (Arsiero et al. 2007) and slow afterhyperpolarization (sAHP) currents (Higgs et al. 2006), but the manner in which they were related to increased fluctuation sensitivity was unclear. For example, in Higgs et al. (2006) it was found that firing rates were more sensitive to input fluctuations in rat pyramidal neurons with large sAHP currents compared to those with small sAHP currents. With the application of α-methyl-5-HT, a 5-HT$_2$ agonist that reduces sAHP currents, the sensitivity of the neurons to input fluctuations was reduced. When artificial sAHP currents were enhanced using dynamic clamp, the sensitivity of firing rates to input fluctuations was increased.

The present results more directly implicate adaptive mechanisms that have a voltage-dependent time constant that is large for peri-threshold values. However, in higher dimensional neurons it is possible that multiple time scales of currents active at peri-threshold voltages may interact such that slow currents with voltage-independent time constants also enhance fluctuation sensitivity. Certainly, in 3D models slow adaptive currents with voltage-independent time constants can alter the geometry of the *V*-nullcline such that trajectories are slower near threshold, although this has a weaker effect than a voltage-dependent recovery time constant.

In summary, we have categorized neurons into Types A, B+, and B− according to how the mean firing rate varies with input mean and variance. We identify three ways that a neuron might increase its sensitivity to input fluctuations: (1) by decreasing its effective sodium pool, or otherwise decreasing the inward-to-outward current ratio prior to spiking, so that the neuron's rest state is always stable (Lundstrom et al. 2008), (2) by changing conductances, or other parameters, to alter the shape of the *V*-nullcline, especially the width between the left and center arms that determines the barrier height ΔU, and (3) by slowing a "recovery" time constant such that the neuron's voltage spends a relatively long time near but below threshold. We believe that this analysis provides a simple and general picture of how the neuron's computation in terms of its *f–I* curves depends on intrinsic parameters. These results help us better understand what from a biophysical standpoint determines or alters a neuron's input/output properties and may assist in predicting the functional effect of seemingly minor homeostatic or adaptive sub-cellular processes.

Acknowledgments We thank Bard Ermentrout for helpful conversations during initial stages of the project at the Marine Biological Laboratory's Methods in Computational Neuroscience 2007 course, Matthew Higgs and Michele Giugliano for helpful discussions and providing data for a figure, and Randy Powers and Sungho Hong for comments on a draft of this manuscript.

Funding This work was supported by a Burroughs-Wellcome Careers at the Scientific Interface grant and a McKnight Scholar Award; BNL was supported by grant number F30NS055650 from the National Institute of Neurological Disorders and Stroke, the Medical Scientist Training Program at UW supported by the National Institute of General Medical Sciences, and an ARCS fellowship; WJS was supported by a VA Merit Review.

Author contributions Conceived of, designed, and performed the simulations: BL. Analyzed the data: BL MF LS AF. Wrote the paper: BL MF WS AF. Developed the conceptual framework: BL MF LS WS AF.

References

Arfken, G. B., & Weber, H. -J. (1995). *Mathematical methods for physicists* (4th ed.). San Diego: Academic.

Arsiero, M., Luscher, H. R., Lundstrom, B. N., & Giugliano, M. (2007). The impact of input fluctuations on the frequency-current relationships of layer 5 pyramidal neurons in the rat medial prefrontal cortex. *The Journal of Neuroscience, 27*, 3274–3284. doi:10.1523/JNEUROSCI.4937-06.2007.

Benda, J., Longtin, A., & Maler, L. (2005). Spike-frequency adaptation separates transient communication signals from background oscillations. *The Journal of Neuroscience, 25*, 2312–2321. doi:10.1523/JNEUROSCI.4795-04.2005.

Chance, F. S., Abbott, L. F., & Reyes, A. D. (2002). Gain modulation from background synaptic input. *Neuron, 35*, 773–782. doi:10.1016/S0896-6273(02)00820-6.

Connor, J. A., & Stevens, C. F. (1971). Prediction of repetitive firing behaviour from voltage clamp data on an isolated neurone soma. *The Journal of Physiology, 213*, 31–53.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience : Computational and mathematical modeling of neural systems*. Cambridge, MA: Massachusetts Institute of Technology Press.

Destexhe, A., Rudolph, M., Fellous, J. M., & Sejnowski, T. J. (2001). Fluctuating synaptic conductances recreate *in vivo*-like activity in neocortical neurons. *Neuroscience, 107*, 13–24. doi:10.1016/S0306-4522(01)00344-X.

Destexhe, A., Rudolph, M., & Pare, D. (2003). The high-conductance state of neocortical neurons *in vivo*. *Nature Reviews. Neuroscience, 4*, 739–751. doi:10.1038/nrn1198.

DeVille, R. E., Vanden-Eijnden, E., & Muratov, C. B. (2005). Two distinct mechanisms of coherence in randomly perturbed dynamical systems. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, 72*, 031105. doi:10.1103/PhysRevE.72.031105.

Ermentrout, B. (1998). Linearization of F–I curves by adaptation. *Neural Computation, 10*, 1721–1729. doi:10.1162/089976698300017106.

Fairhall, A. L., Lewen, G. D., Bialek, W., & de Ruyter Van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature, 412*, 787–792. doi:10.1038/35090500.

Fellous, J. M., Rudolph, M., Destexhe, A., & Sejnowski, T. J. (2003). Synaptic background noise controls the input/output characteristics of single cells in an *in vitro* model of *in vivo* activity. *Neuroscience, 122*, 811–829. doi:10.1016/j.neuroscience.2003.08.027.

Fleidervish, I. A., Friedman, A., & Gutnick, M. J. (1996). Slow inactivation of Na+ current and slow cumulative spike adaptation in mouse and guinea-pig neocortical neurones in slices. *The Journal of Physiology, 493*(Pt 1), 83–97.

Gerstner, W., & Kistler, W. M. (2002). *Spiking neuron models : Single neurons, populations, plasticity*. Cambridge, UK: Cambridge University Press.

Gutkin, B. S., & Ermentrout, G. B. (1998). Dynamics of membrane excitability determine interspike interval variability: a link between spike generation mechanisms and cortical spike train statistics. *Neural Computation, 10*, 1047–1065. doi:10.1162/089976698300017331.

Higgs, M. H., Slee, S. J., & Spain, W. J. (2006). Diversity of gain modulation by noise in neocortical neurons: regulation by the slow afterhyperpolarization conductance. *The Journal of Neuroscience, 26*, 8787–8799. doi:10.1523/JNEUROSCI.1792-06.2006.

Hodgkin, A. L. (1948). The local electric changes associated with repetitive action in a non-medullated axon. *The Journal of Physiology, 107*, 165–181.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology, 117*, 500–544.

Hong, S., Aguera y Arcas, B., & Fairhall, A. L. (2007). Single neuron computation: from dynamical system to feature detector. *Neural Computation, 19*, 3133–3172. doi:10.1162/neco.2007.19.12.3133.

Hong, S., Lundstrom, B. N., & Fairhall, A. L. (2008). Intrinsic gain modulation and adaptive neural coding. *PLoS Computational Biology, 4*, e1000119. doi:10.1371/journal.pcbi.1000119.

Izhikevich, E. M. (2007). *Dynamical systems in neuroscience : The geometry of excitability and bursting*. Cambridge, MA: MIT Press.

Koch, C. (1999). *Biophysics of computation: information processing in single neurons*. New York: Oxford University Press.

Konig, P., Engel, A. K., & Singer, W. (1996). Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends in Neurosciences, 19*, 130–137. doi:10.1016/S0166-2236(96)80019-1.

Lundstrom, B. N., & Fairhall, A. L. (2006). Decoding stimulus variance from a distributional neural code of interspike intervals. *The Journal of Neuroscience, 26*, 9030–9037. doi:10.1523/JNEUROSCI.0225-06.2006.

Lundstrom, B. N., Hong, S., Higgs, M. H., & Fairhall, A. L. (2008). Two computational regimes of a single-compartment neuron separated by a planar boundary in conductance space. Neural Comput, 20, 1239–1260.

Moreno, R., de la Rocha, J., Renart, A., & Parga, N. (2002). Response of spiking neurons to correlated inputs. *Physical Review Letters, 89*, 288101. doi:10.1103/PhysRevLett.89.288101.

Morris, C., & Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal, 35*, 193–213. doi:10.1016/S0006-3495(81)84782-0.

Prescott, S. A., Ratte, S., De Koninck, Y., & Sejnowski, T. J. (2006). Nonlinear interaction between shunting and adaptation controls a switch between integration and coincidence detection in pyramidal neurons. *The Journal of Neuroscience*, *26*, 9084–9097. doi:10.1523/JNEUROSCI.1388-06.2006.

Rauch, A., La Camera, G., Luscher, H. R., Senn, W., & Fusi, S. (2003). Neocortical pyramidal cells respond as integrate-and-fire neurons to *in vivo*-like input currents. *Journal of Neurophysiology*, *90*, 1598–1612. doi:10.1152/jn.00293.2003.

Richardson, M. J. (2004). Effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *69*, 051918. doi:10.1103/PhysRevE.69.051918.

Rinzel, J., & Ermentrout, B. (1998). *Analysis of neural excitability and oscillations*. In C. Koch, I. Segev, & (Eds.), (pp. 251–291, 2nd ed.). Cambridge, Massachusetts: MIT Press.

Robinson, H. P., & Harsch, A. (2002). Stages of spike time variability during neuronal responses to transient inputs. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *66*, 061902. doi:10.1103/PhysRevE.66.061902.

Rudolph, M., & Destexhe, A. (2005). An extended analytic expression for the membrane potential distribution of conductance-based synaptic noise. *Neural Computation*, *17*, 2301–2315. doi:10.1162/0899766054796932.

Rudolph, M., & Destexhe, A. (2006). On the use of analytical expressions for the voltage distribution to analyze intracellular recordings. *Neural Computation*, *18*, 2917–2922. doi:10.1162/neco.2006.18.12.2917.

Rush, M. E., & Rinzel, J. (1995). The potassium A-current, low firing rates and rebound excitation in Hodgkin–Huxley models. *Bulletin of Mathematical Biology*, *57*, 899–929.

Shadlen, M. N., & Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, *4*, 569–579. doi:10.1016/0959-4388(94)90059-0.

Slee, S. J., Higgs, M. H., Fairhall, A. L., & Spain, W. J. (2005). Two-dimensional time coding in the auditory brainstem. *The Journal of Neuroscience*, *25*, 9978–9988. doi:10.1523/JNEUROSCI.2666-05.2005.

Strogatz, S. H. (1994). *Nonlinear dynamics and Chaos: with applications to physics, biology, chemistry, and engineering*. Reading, Mass.: Addison-Wesley.

Tateno, T., & Pakdaman, K. (2004). Random dynamics of the Morris–Lecar neural model. *Chaos (Woodbury, N.Y.)*, *14*, 511–530. doi: doi:10.1063/1.1756118.

VanRullen, R., Guyonneau, R., & Thorpe, S. J. (2005). Spike times make sense. *Trends in Neurosciences*, *28*, 1–4. doi:10.1016/j.tins.2004.10.010.

ARTICLE Communicated by Paul Bressloff

Computation in a Single Neuron: Hodgkin and Huxley Revisited

Blaise Agüera y Arcas
blaisea@princeton.edu
Rare Books Library, Princeton University, Princeton, NJ 08544, U.S.A.

Adrienne L. Fairhall
fairhall@princeton.edu
NEC Research Institute, Princeton, NJ 08540, and Department of Molecular Biology, Princeton, NJ 08544, U.S.A.

William Bialek
wbialek@princeton.edu
NEC Research Institute, Princeton, NJ 08540, and Department of Physics, Princeton, NJ 08544, U.S.A.

A spiking neuron "computes" by transforming a complex dynamical input into a train of action potentials, or spikes. The computation performed by the neuron can be formulated as dimensional reduction, or feature detection, followed by a nonlinear decision function over the low-dimensional space. Generalizations of the reverse correlation technique with white noise input provide a numerical strategy for extracting the relevant low-dimensional features from experimental data, and information theory can be used to evaluate the quality of the low–dimensional approximation. We apply these methods to analyze the simplest biophysically realistic model neuron, the Hodgkin–Huxley (HH) model, using this system to illustrate the general methodological issues. We focus on the features in the stimulus that trigger a spike, explicitly eliminating the effects of interactions between spikes. One can approximate this triggering "feature space" as a two-dimensional linear subspace in the high-dimensional space of input histories, capturing in this way a substantial fraction of the mutual information between inputs and spike time. We find that an even better approximation, however, is to describe the relevant subspace as two dimensional but curved; in this way, we can capture 90% of the mutual information even at high time resolution. Our analysis provides a new understanding of the computational properties of the HH model. While it is common to approximate neural behavior as "integrate and fire," the HH model is not an integrator nor is it well described by a single threshold.

Neural Computation **15, 1715–1749** (2003) © 2003 Massachusetts Institute of Technology

1 Introduction

On short timescales, one can conceive of a single neuron as a computational device that maps inputs at its synapses into a sequence of action potentials or spikes. To a good approximation, the dynamics of this mapping are determined by the kinetic properties of ion channels in the neuron's membrane. In the 50 years since the pioneering work of Hodgkin and Huxley, we have seen the evolution of an ever more detailed description of channel kinetics, making it plausible that the short time dynamics of almost any neuron we encounter will be understandable in terms of interactions among a mixture of diverse but known channel types (Hille, 1992; Koch, 1999). The existence of so nearly complete a microscopic picture of single-neuron dynamics brings into focus a very different question: What does the neuron compute? Although models in the Hodgkin–Huxley (HH) tradition define a dynamical system that will reproduce the behavior of the neuron, this description in terms of differential equations is far from our intuition about—or the formal description of—computation.

The problem of what neurons compute is one instance of a more general problem in modern quantitative biology and biophysics: Given a progressively more complete microscopic description of proteins and their interactions, how do we understand the emergence of function? In the case of neurons, the proteins are the ion channels, and the interactions are very simple: current flows through open channels, charging the cell's capacitance, and all channels experience the resulting voltage. Arguably, there is no other network of interacting proteins for which the relevant equations are known in such detail; indeed, some efforts to understand function and computation in other networks of proteins make use of analogies to neural systems (Bray, 1995). Despite the relative completeness of our microscopic picture for neurons, there remains a huge gap between the description of molecular kinetics and the understanding of function. Given some complex dynamic input to a neuron, we might be able to simulate the spike train that will result, but we are hard pressed to look at the equations for channel kinetics and say that this transformation from inputs to spikes is equivalent to some simple (or perhaps not so simple) computation such as filtering, thresholding, coincidence detection, or feature extraction.

Perhaps the problem of understanding computational function in a model of ion channel dynamics is a symptom of a much deeper mathematical difficulty. Despite the fact that all computers are dynamical systems, the natural mathematical objects in dynamical systems theory are very different from those in the theory of computation, and it is not clear how to connect these different formal schemes. Finding a general mapping from dynamical systems to their equivalent computational functions is a grand challenge, but we will take a more modest approach.

We believe that a key intuition for understanding neural computation is the concept of feature selectivity: while the space of inputs to a neuron—

whether we think of inputs as arriving at the synapses or being driven by sensory signals outside the brain—is vast, individual neurons are sensitive only to some restricted set of features in this vast space. The most general way to formalize this intuition is to say that we can compress (in the information-theoretic sense) our description of the inputs without losing any information about the neural output (Tishby, Pereira, & Bialek, 1999). We might hope that this selective compression of the input data has a simple geometric description, so that the relevant bits about the input correspond to coordinates along some restricted set of relevant dimensions in the space of inputs. If this is the case, feature selectivity should be formalized as a reduction of dimensionality (de Ruyter van Steveninck & Bialek, 1988), and this is the approach we follow here. Closely related work on the use of dimensionality reduction to analyze neural feature selectivity has been described in recent work (Bialek & de Ruyter van Steveninck, 2003; Sharpee, Rust, & Bialek, in press).

Here, we develop the idea of dimensionality reduction as a tool for analysis of neural computation and apply these tools to the HH model. While our initial goal was to test new analysis methods in the context of a presumably simple and well-understood model, we have found that the HH neuron performs a computation of surprising richness. Preliminary accounts of these results have already appeared (Agüera y Arcas, 1998; Agüera y Arcas, Bialek, & Fairhall, 2001).

2 Dimensionality Reduction

Neurons take input signals at their synapses and give as output sequences of spikes. To characterize a neuron completely is to identify the mapping between neuronal input and the spike train the neuron produces in response. In the absence of any simplifying assumptions, this requires probing the system with every possible input. Most often, these inputs are spikes from other neurons; each neuron typically has of order $N \sim 10^3$ presynaptic connections. If the system operates at 1 msec resolution and the time window of relevant inputs is 40 msec, then we can think of a single neuron as having an input described by a $\sim 4 \times 10^4$ bit word—the presence or absence of a spike in each 1 msec bin for each presynaptic cell—which is then mapped to a one (spike) or zero (no spike). More realistically, if average spike rates are $\sim 10\ \text{s}^{-1}$, the input words can be compressed by a factor of 10. In this picture, a neuron computes a Boolean function over roughly 4000 variables. Clearly one cannot sample every one of the $\sim 2^{4000}$ inputs to identify the neural computation. Progress requires making some simplifying assumption about the function computed by the neuron so that we can vastly reduce the space of possibilities over which to search. We use the idea of dimensionality reduction in this spirit, as a simplifying assumption that allows us to make progress but that also must be tested directly.

The ideas of feature selectivity and dimensionality reduction have a long history in neurobiology. The idea of receptive fields as formulated by Hartline, Kuffler, and Barlow for the visual system gave a picture of neurons as having a template against which images would be correlated (Hartline, 1940; Kuffler, 1953; Barlow, 1953). If we think of images as vectors in a high-dimensional space, with coordinates determined by the intensities of each pixel, then the simplest receptive field models describe the neuron as sensitive to only one direction or projection in this high-dimensional space. This picture of projection followed by thresholding or some other nonlinearity to determine the probability of spike generation was formalized in the linear perceptron (Rosenblatt, 1958, 1962). In subsequent work, Barlow, Hill, and Levick (1964) characterized neurons in which the receptive field has subregions in space and time such that summation is at least approximately linear in each subregion but these summed signals interact nonlinearly, for example, to generate direction selectivity and motion sensitivity. We can think of Hubel and Wiesel's description of complex and hypercomplex cells (Hubel & Wiesel, 1962) again as a picture of approximately linear summation within subregions followed by nonlinear operations on these multiple summed signals. More formally, the proper combination of linear summation and nonlinear or logical operations may provide a useful bridge from receptive field properties to proper geometric primitives in visual computation (Iverson & Zucker, 1995). In the same way that a single receptive field or perceptron model has one relevant dimension in the space of visual stimuli, these more complex cells have as many relevant dimensions as there are independent subregions of the receptive field. Although this number is larger than one, it still is much smaller than the full dimensionality of the possible spatiotemporal variations in visual inputs.

The idea that neurons in the auditory system might be described by a filter followed by a nonlinear transformation to determine the probability of spike generation was the inspiration for de Boer's development (de Boer & Kuyper, 1968) of triggered or reverse correlation. Modern uses of reverse correlation to characterize the filtering or receptive field properties of a neuron often emphasize that this approach provides a "linear approximation" to the input-output properties of the cell, but the original idea was almost the opposite: neurons clearly are nonlinear devices, but this is separate from the question of whether the probability of generating a spike is determined by a simple projection of the sensory input onto a single filter or template. In fact, as explained by Rieke, Warland, Bialek, and de Ruyter van Steveninck (1997), linearity is seldom a good approximation for the neural input-output relation, but if there is one relevant dimension, then (provided that input signals are chosen with suitable statistics) the reverse correlation method is guaranteed to find this one special direction in the space of inputs to which the neuron is sensitive. While the reverse correlation method is guaranteed to find the one relevant dimension if it exists, the method does not include any way of testing for other relevant dimensions, or more generally for measuring the dimensionality of the relevant subspace.

The idea of characterizing neural responses directly as the reduction of dimensionality emerged from studies (de Ruyter van Steveninck & Bialek, 1988) of a motion-sensitive neuron in the fly visual system. In particular, this work suggested that it is possible to estimate the dimensionality of the relevant subspace rather than just assuming that it is small (or equal to one). More recent work on the fly visual system has exploited the idea of dimensionality reduction to probe both the structure and adaptation of the neural code (Brenner, Bialek, & de Ruyter van Steveninck, 2000; Fairhall, Lewen, Bialek, & de Ruyter van Steveninck, 2001) and the nature of the computation that extracts the motion signal from the spatiotemporal array of photoreceptor inputs (Bialek & de Ruyter van Steveninck, 2003). Here we review the ideas of dimensionality reduction from previous work; extensions of these ideas begin in section 3.

In the spirit of neural network models, we will simplify away the spatial structure of neurons and consider time-dependent currents $I(t)$ injected into a point–like neuron. While this misses much of the complexity of real cells, we will find that even this system is highly nontrivial. If the input is an injected current, then the neuron maps the history of this current, $I(t < t_0)$, into the presence or absence of a spike at time t_0. More generally, we might imagine that the cell (or our description) is noisy, so that there is a probability of spiking $P[\text{spike at } t_0 \mid I(t < t_0)]$ that depends on the current history. The dependence on the history of the current means that the input signal still is high dimensional, even without spatial dependence. Working at time resolution Δt and assuming that currents in a window of size T are relevant to the decision to spike, the input space is of dimension $D = T/\Delta t$, where D is often of order 100.

The idea of dimensionality reduction is that the probability of spike generation is sensitive only to some limited number of dimensions K within the D-dimensional space of inputs. We begin our analysis by searching for linear subspaces, that is, a set of signals $s_1, s_2, \ldots, s_K$ that can be constructed by filtering the current,

$$s_\mu = \int_0^\infty dt f_\mu(t) I(t_0 - t), \tag{2.1}$$

so that the probability of spiking depends on only this small set of signals,

$$P[\text{spike at } t_0 \mid I(t < t_0)] = P[\text{spike at } t_0] g(s_1, s_2, \ldots, s_K), \tag{2.2}$$

where the inclusion of the average probability of spiking, $P[\text{spike at } t_0]$, leaves g dimensionless. If we think of the current $I(t_0 - T < t < t_0)$ as a D-dimensional vector, with one dimension for each discrete sample at spacing Δt, then the filtered signals s_i are linear projections of this vector. In this formulation, characterizing the computation done by a neuron involves three steps:

1. Estimate the number of relevant stimulus dimensions K, with the hope that there will be many fewer than the original dimensionality D.
2. Identify a set of filters that project into this relevant subspace.
3. Characterize the nonlinear function $g(\vec{s})$.

The classical perceptron–like cell of neural network theory would have only one relevant dimension, given by the vector of weights, and a simple form for g, typically a sigmoid.

Rather than trying to look directly at the distribution of spikes given stimuli, we follow de Ruyter van Steveninck and Bialek (1988) and consider the distribution of signals conditional on the response, $P[I(t < t_0) \mid \text{spike at } t_0]$, also called the response conditional ensemble (RCE); these are related by Bayes' rule,

$$\frac{P[\text{spike at } t_0 \mid I(t < t_0)]}{P[\text{spike at } t_0]} = \frac{P[I(t < t_0) \mid \text{spike at } t_0]}{P[I(t < t_0)]}. \tag{2.3}$$

We can now compute various moments of the RCE. The first moment is the spike-triggered average stimulus (STA),

$$\text{STA}(\tau) = \int [dI] P[I(t < t_0) \mid \text{spike at } t_0] I(t_0 - \tau), \tag{2.4}$$

which is the object that one computes in reverse correlation (de Boer & Kuyper, 1968; Rieke et al., 1997). If we choose the distribution of input stimuli $P[I(t < t_0)]$ to be gaussian white noise, then for a perceptron–like neuron sensitive to only one direction in stimulus space, it can be shown that the STA or first moment of the RCE is proportional to the vector or filter $f(\tau)$ that defines this direction (Rieke et al., 1997).

Although it is a theorem that the STA is proportional to the relevant filter $f(\tau)$, in principle it is possible that the proportionality constant is zero, most plausibly if the neuron's response has some symmetry, such as phase invariance in the response of high-frequency auditory neurons. It also is worth noting that what is really important in this analysis is the gaussian distribution of the stimuli, not the "whiteness" of the spectrum. For nonwhite but gaussian inputs, the STA measures the relevant filter blurred by the correlation function of the inputs, and hence the true filter can be recovered (at least in principle) by deconvolution. For nongaussian signals and nonlinear neurons, there is no corresponding guarantee that the selectivity of the neuron can be separated from correlations in the stimulus (Sharpee et al., in press).

To obtain more than one relevant direction (or to reveal relevant directions when symmetries cause the STA to vanish), we proceed to second order and compute the covariance matrix of fluctuations around the spike-triggered average,

$$C_{\text{spike}}(\tau, \tau') = \int [dI] P[I(t < t_0) \mid \text{spike at } t_0] I(t_0 - \tau) I(t_0 - \tau') - \text{STA}(\tau)\text{STA}(\tau'). \tag{2.5}$$

In the same way that we compare the spike-triggered average to some constant average level of the signal in the whole experiment, we compare the covariance matrix C_{spike} with the covariance of the signal averaged over the whole experiment,

$$C_{\text{prior}}(\tau, \tau') = \int [dI] P[I(t < t_0)] I(t_0 - \tau) I(t_0 - \tau'), \tag{2.6}$$

to construct the change in the covariance matrix,

$$\Delta C = C_{\text{spike}} - C_{\text{prior}}. \tag{2.7}$$

With time resolution Δt in a window of duration T as above, all of these covariances are $D \times D$ matrices. In the same way that the spike-triggered average has the clearest interpretation when we choose inputs from a gaussian distribution, ΔC also has the clearest interpretation in this case. Specifically, if inputs are drawn from a gaussian distribution, then it can be shown that (Bialek & de Ruyter van Steveninck, 2003):

1. If the neuron is sensitive to a limited set of K-input dimensions as in equation 2.2, then ΔC will have only K nonzero eigenvalues.[1] In this way, we can measure directly the dimensionality K of the relevant subspace.
2. If the distribution of inputs is both gaussian and white, then the eigenvectors associated with the nonzero eigenvalues span the same space as that spanned by the filters $\{f_\mu(\tau)\}$.
3. For nonwhite (correlated) but still gaussian inputs, the eigenvectors span the space of the filters $\{f_\mu(\tau)\}$ blurred by convolution with the correlation function of the inputs.

Thus, the analysis of ΔC for neurons responding to gaussian inputs should allow us to identify the subspace of inputs of relevance and test specifically the hypothesis that this subspace is of low dimension.

---

[1] As with the STA, it is in principle possible that symmetries or accidental features of the function $g(\vec{s})$ would cause some of the K eigenvalues to vanish, but this is very unlikely.

Several points are worth noting. First, except in special cases, the eigenvectors of ΔC and the filters $\{f_\mu(\tau)\}$ are not the principal components of the RCE, and hence this analysis of ΔC is not a principal component analysis. Second, the nonzero eigenvalues of ΔC can be either positive or negative, depending on whether the variance of inputs along that particular direction is larger or smaller in the neighborhood of a spike. Third, although the eigenvectors span the relevant subspace, these eigenvectors do not form a preferred coordinate system within this subspace. Finally, we emphasize that dimensionality reduction—identification of the relevant subspace—is only the first step in our analysis of the computation done by a neuron.

3 Measuring the Success of Dimensionality Reduction

The claim that certain stimulus features are most relevant is in effect a model for the neuron, so the next question is how to measure the effectiveness or accuracy of this model. Several different ideas have been suggested in the literature as ways of testing models based on linear receptive fields in the visual system (Stanley, Lei, & Dan, 1999; Keat, Reinagel, Reid, & Meister, 2001) or linear spectrotemporal receptive fields in the auditory system (Theunissen, Sen, & Doupe, 2000). These methods have in common that they introduce a metric to measure performance—for example, mean square error in predicting the firing rate as averaged over some window of time. Ideally, we would like to have a performance measure that avoids any arbitrariness in the choice of metric, and such metric-free measures are provided uniquely by information theory (Shannon, 1948; Cover & Thomas, 1991).

Observing the arrival time t_0 of a single spike provides a certain amount of information about the input signals. Since information is mutual, we can also say that knowing the input signal trajectory $I(t < t_0)$ provides information about the arrival time of the spike. If "details are irrelevant," then we should be able to discard these details from our description of the stimulus and yet preserve the mutual information between the stimulus and spike arrival times (for an abstract discussion of such selective compression, see Tishby et al., 1999). In constructing our low-dimensional model, we represent the complete (D-dimensional) stimulus $I(t < t_0)$ by a smaller number ($K < D$) of dimensions $\vec{s} = (s_1, s_2, \ldots, s_K)$.

The mutual information $\mathcal{I}[I(t < t_0); t_0]$ is a property of the neuron itself, while the mutual information $\mathcal{I}[\vec{s}; t_0]$ characterizes how much our reduced description of the stimulus can tell us about when spikes will occur. Necessarily, our reduction of dimensionality causes a loss of information, so that

$$\mathcal{I}[\vec{s}; t_0] \leq \mathcal{I}[I(t < t_0); t_0], \tag{3.1}$$

but if our reduced description really captures the computation done by the neuron, then the two information measures will be very close. In particular, if the neuron were described exactly by a lower-dimensional model—as for a linear perceptron or for an integrate-and-fire neuron (Agüera y Arcas & Fairhall, 2003)—then the two information measures would be equal. More generally, the ratio $\mathcal{I}[\vec{s}; t_0]/\mathcal{I}[I(t < t_0); t_0]$ quantifies the efficiency of the low-dimensional model, measuring the fraction of information about spike arrival times that our K dimensions capture from the full signal $I(t < t_0)$.

As shown by Brenner, Strong, Koberle, Bialek, and de Ruyter van Steveninck (2000), the arrival time of a single spike provides an information,

$$\mathcal{I}[I(t < t_0); t_0] \equiv I_{\text{one spike}} = \frac{1}{T}\int_0^T dt \frac{r(t)}{\bar{r}} \log_2\left[\frac{r(t)}{\bar{r}}\right], \tag{3.2}$$

where $r(t)$ is the time-dependent spike rate, $\bar{r}$ is the average spike rate, and $\langle \cdots \rangle$ denotes an average over time. In principle, information should be calculated as an average over the distribution of stimuli, but the ergodicity of the stimulus justifies replacing this ensemble average with a time average. For a deterministic system like the HH equations, the spike rate is a singular function of time: given the inputs $I(t)$, spikes occur at definite times with no randomness or irreproducibility. If we observe these responses with a time resolution Δt, then for Δt sufficiently small, the rate $r(t)$ at any time t either is zero or corresponds to a single spike occurring in one bin of size Δt, that is, $r = 1/\Delta t$. Thus, the information carried by a single spike is

$$I_{\text{one spike}} = -\log_2 \bar{r}\Delta t. \tag{3.3}$$

On the other hand, if the probability of spiking really depends on only the stimulus dimensions $s_1, s_2, \ldots, s_K$, we can substitute

$$\frac{r(t)}{\bar{r}} \rightarrow \frac{P(\vec{s} \mid \text{spike at } t)}{P(\vec{s})}. \tag{3.4}$$

Replacing the time averages in equation 3.2 with ensemble averages, we find

$$\mathcal{I}[\vec{s}; t_0] \equiv I^{\vec{s}}_{\text{one spike}} = \int d^K s P(\vec{s} \mid \text{spike at } t) \log_2\left[\frac{P(\vec{s} \mid \text{spike at } t)}{P(\vec{s})}\right] \tag{3.5}$$

(for details of these arguments, see Brenner, Strong, et al., 2000). This allows us to compare the information captured by the K-dimensional reduced model with the true information carried by single spikes in the spike train.

For reasons that we will discuss in the following section, and as was pointed out in Agüera y Arcas et al. (2001) and Agüera y Arcas and Fairhall (2003), we will be considering isolated spikes—those separated from previous spikes by a period of silence. This has important consequences for our analysis. Most significantly, as we will be considering spikes that occur on a background of silence, the relevant stimulus ensemble, conditioned on the silence, is no longer gaussian. Further, we will need to refine our information estimate.

The derivation of equation 3.2 makes clear that a similar formula must determine the information carried by the occurrence time of any event, not just single spikes; we can define an event rate in place of the spike rate and then calculate the information carried by these events (Brenner, Strong, et al., 2000). In the case here, we wish to compute the information obtained by observing an isolated spike, or equivalently by the event silence+spike. This is straightforward: we replace the spike rate by the rate of isolated spikes, and equation 3.2 will give us the information carried by the arrival time of a single isolated spike. The problem is that this information includes both the information carried by the occurrence of the spike and the information conveyed in the condition that there were no spikes in the preceding t_{silence} msec (for an early discussion of the information carried by silence, see de Ruyter van Steveninck & Bialek, 1988). We would like to separate these contributions, since our idea of dimensionality reduction applies only to the triggering of a spike, not to the temporally extended condition of nonspiking.

To separate the information carried by the isolated spike itself, we have to ask how much information we gain by seeing an isolated spike given that the condition for isolation has already been met. As discussed by Brenner, Strong, et al. (2000), we can compute this information by thinking about the distribution of times at which the isolated spike can occur. Given that we know the input stimulus, the distribution of times at which a single isolated spike will be observed is proportional to $r_{\text{iso}}(t)$, the time-dependent rate or peristimulus time histogram for isolated spikes. With proper normalization, we have

$$P_{\text{iso}}(t \mid \text{inputs}) = \frac{1}{T} \cdot \frac{1}{\bar{r}_{\text{iso}}} r_{\text{iso}}(t), \tag{3.6}$$

where T is duration of the (long) window in which we can look for the spike and $\bar{r}_{\text{iso}}$ is the average rate of isolated spikes. This distribution has an entropy,

$$\begin{aligned} S_{\text{iso}}(t \mid \text{inputs}) &= -\int_0^T dt\, P_{\text{iso}}(t \mid \text{inputs}) \log_2 P_{\text{iso}}(t \mid \text{inputs}) && (3.7) \\ &= -\frac{1}{T}\int_0^T dt \frac{r_{\text{iso}}(t)}{\bar{r}_{\text{iso}}} \log_2 \left[\frac{1}{T} \cdot \frac{r_{\text{iso}}(t)}{\bar{r}_{\text{iso}}}\right] && (3.8) \\ &= \log_2(T\bar{r}_{\text{iso}}\Delta t) \text{ bits}, && (3.9) \end{aligned}$$

where again we use the fact that for a deterministic system, the time-dependent rate must be either zero or the maximum allowed by our time resolution Δt. To compute the information carried by a single spike, we need to compare this entropy with the total entropy possible when we do not know the inputs.

It is tempting to think that without knowledge of the inputs, an isolated spike is equally likely to occur anywhere in the window of size T, which leads us back to equation 3.3 with $\bar{r}$ replaced by $\bar{r}_{\text{iso}}$. In this case, however, we are assuming that the condition for isolation has already been met. Thus, even without observing the inputs, we know that isolated spikes can occur only in windows of time whose total length is $T_{\text{silence}} = T \cdot P_{\text{silence}}$, where P_{silence} is the probability that any moment in time is at least t_{silence} after the most recent spike. Thus, the total entropy of isolated spike arrival times (given that the condition for silence has been met) is reduced from $\log_2 T$ to

$$S_{\text{iso}}(t \mid \text{silence}) = \log_2(T \cdot P_{\text{silence}}), \tag{3.10}$$

and the information that the spike carries beyond what we know from the silence itself is

$$\Delta I_{\text{iso spike}} = S_{\text{iso}}(t \mid \text{silence}) - S_{\text{iso}}(t \mid \text{inputs}) \tag{3.11}$$

$$= \frac{1}{T}\int_0^T dt \frac{r_{\text{iso}}(t)}{\bar{r}_{\text{iso}}} \log_2\left[\frac{r_{\text{iso}}(t)}{\bar{r}_{\text{iso}}} \cdot P_{\text{silence}}\right] \tag{3.12}$$

$$= -\log_2(\bar{r}_{\text{iso}}\Delta t) + \log_2 P_{\text{silence}} \text{ bits}. \tag{3.13}$$

This information, which is defined independent of any model for the feature selectivity of the neuron, provides the benchmark against which our reduction of dimensionality will be measured. To make the comparison, however, we need the analog of equation 3.5.

Equation 3.12 provides us with an expression for the information conveyed by isolated spikes in terms of the probability that these spikes occur at particular times; this is analogous to equation 3.2 for single (nonisolated) spikes. If we follow a path analogous to that which leads from equation 3.2 to equation 3.5, we find an expression for the information that an isolated spike provides about the K stimulus dimensions $\vec{s}$:

$$\Delta I^{\vec{s}}_{\text{iso spike}} = \int d\vec{s}\, P(\vec{s} \mid \text{iso spike at } t) \log_2\left[\frac{P(\vec{s} \mid \text{iso spike at } t)}{P(\vec{s} \mid \text{silence})}\right] + \langle \log_2 P(\text{silence} \mid \vec{s}) \rangle, \tag{3.14}$$

where the prior is now also conditioned on silence: $P(\vec{s} \mid \text{silence})$ is the distribution of $\vec{s}$ given that $\vec{s}$ is preceded by a silence of at least t_{silence}. Notice that this silence-conditioned distribution is not knowable a priori, and in particular it is not gaussian; $P(\vec{s} \mid \text{silence})$ must be sampled from data.

The last term in equation 3.14 is the entropy of a binary variable that indicates whether particular moments in time are silent given knowledge of the stimulus. Again, since the HH model is deterministic, this conditional entropy should be zero if we keep a complete description of the stimulus. In fact, we are not interested in describing those features of the stimulus that lead to silence, and it is not fair (as we will see) to judge the success of dimensionality reduction by looking at the prediction of silence, which necessarily involves multiple dimensions. To make a meaningful comparison, then, we will assume that there is a perfect description of the stimulus conditions leading to silence and focus on the stimulus features that trigger the isolated spike. When we approximate these features by the K-dimensional space $\vec{s}$, we capture an amount of information,

$$\Delta I^{\vec{s}}_{\text{iso spike}} = \int d\vec{s}\, P(\vec{s} \mid \text{iso spike at } t) \log_2 \left[\frac{P(\vec{s} \mid \text{iso spike at } t)}{P(\vec{s} \mid \text{silence})} \right]. \quad (3.15)$$

This is the information that we can compare with $\Delta I_{\text{iso spike}}$ in equation 3.13 to determine the efficiency of our dimensionality reduction.

4 Characterizing the Hodgkin-Huxley Neuron

For completeness, we begin with a brief review of the dynamics of the space-clamped HH neuron (Hodgkin & Huxley, 1952). Hodgkin and Huxley modeled the dynamics of the current through a patch of membrane with ion-specific conductances:

$$C\frac{dV}{dt} = I(t) - \bar{g}_{\text{K}} n^4 (V - V_{\text{K}}) - \bar{g}_{\text{Na}} m^3 h (V - V_{\text{Na}}) - \bar{g}_l (V - V_l), \quad (4.1)$$

where $I(t)$ is injected current, K and Na subscripts denote potassium– and sodium–related variables, respectively, and l (for "leakage") terms include all other ion conductances with slower dynamics. C is the membrane capacitance. V_{K} and V_{Na} are ion-specific reversal potentials, and V_l is defined such that the total voltage V is exactly zero when the membrane is at rest. $\bar{g}_{\text{K}}$, $\bar{g}_{\text{Na}}$, and $\bar{g}_l$ are empirically determined maximal conductances for the different ion species, and the gating variables n, m, and h (on the interval $[0, 1]$) have their own voltage-dependent dynamics:

$$\begin{aligned} dn/dt &= (0.01V + 0.1)(1 - n)\exp(-0.1V) - 0.125n\exp(V/80), \\ dm/dt &= (0.1V + 2.5)(1 - m)\exp(-0.1V - 1.5) - 4m\exp(V/18), \\ dh/dt &= 0.07(1 - h)\exp(0.05V) - h\exp(-0.1V - 4). \end{aligned} \quad (4.2)$$

We have used the original values for these parameters, except for changing the signs of the voltages to correspond to the modern sign convention: $C = 1$ $\mu\text{F/cm}^2$, $\bar{g}_{\text{K}} = 36\,\text{mS/cm}^2$, $\bar{g}_{\text{Na}} = 120\,\text{mS/cm}^2$, $\bar{g}_l = 0.3\,\text{mS/cm}^2$, $V_{\text{K}} = -12$ mV, $V_{\text{Na}} = +115$ mV, $V_l = +10.613$ mV. We have taken our system to be a $\pi \times 30^2$ μm^2 patch of membrane. We solve these equations numerically using fourth-order Runge–Kutta integration.

The system is driven with a gaussian random noise current $I(t)$, generated by smoothing a gaussian random number stream with an exponential filter to generate a correlation time τ. It is convenient to choose τ to be longer than the time steps of numerical integration, since this guarantees that all functions are smooth on the scale of single time steps. Here, we will always use $\tau = 0.2$ msec, a value that is both less than the timescale over which we discretize the stimulus for analysis and far less than the neuron's capacitative smoothing timescale $RC \sim 3$ msec. $I(t)$ has a standard deviation σ, but since the correlation time is short, the relevant parameter usually is the spectral density $S = \sigma^2\tau$; we also add a DC offset I_0. In the following, we will consider two parameter regimes, $I_0 = 0$ and I_0 a finite value, which leads to more periodic firing.

The integration step size is fixed at 0.05 msec. The key numerical experiments were repeated at a step size of 0.01 msec with identical results. The time of a spike is defined as the moment of maximum voltage, for voltages exceeding a threshold (see Figure 1), estimated to subsample precision by quadratic interpolation. As spikes are both very stereotyped and very large compared to subspiking fluctuations, the precise value of this threshold is unimportant; we have used +20 mV.

4.1 Qualitative Description of Spiking. The first step in our analysis is to use reverse correlation, equation 2.4, to determine the average stimulus feature preceding a spike, the STA. In Figure 1(top), we display the STA in a regime where the spectral density of the input current is 6.5×10^{-4} nA^2 msec. The spike-triggered averages of the gating terms n^4 (proportion of open potassium channels) and m^3h (proportion of open sodium channels) and the membrane voltage V are plotted in Figure 1 (middle and bottom). The error bars mark the standard deviation of the trajectories of these variables.

As expected, the voltage and gating variables follow highly stereotyped trajectories during the ~5 msec surrounding a spike. First, the rapid opening of the sodium channels causes a sharp membrane depolarization (or rise in V); the slower potassium channels then open and repolarize the membrane, leaving it at a slightly lower potential than rest. The potassium channels close gradually, but meanwhile the membrane remains hyperpolarized and, due to its increased permeability to potassium ions, at lower resistance. These effects make it difficult to induce a second spike during this ~15 msec "refractory period." Away from spikes, the resting levels and fluctuations of the voltage and gating variables are quite small. The larger values evident in Figure 1(middle and bottom) by ±15 msec are due to the summed contributions of nearby spikes.

The spike-triggered average current has a largely transient form, so that spikes are on average preceded by an upward swing in current. On the

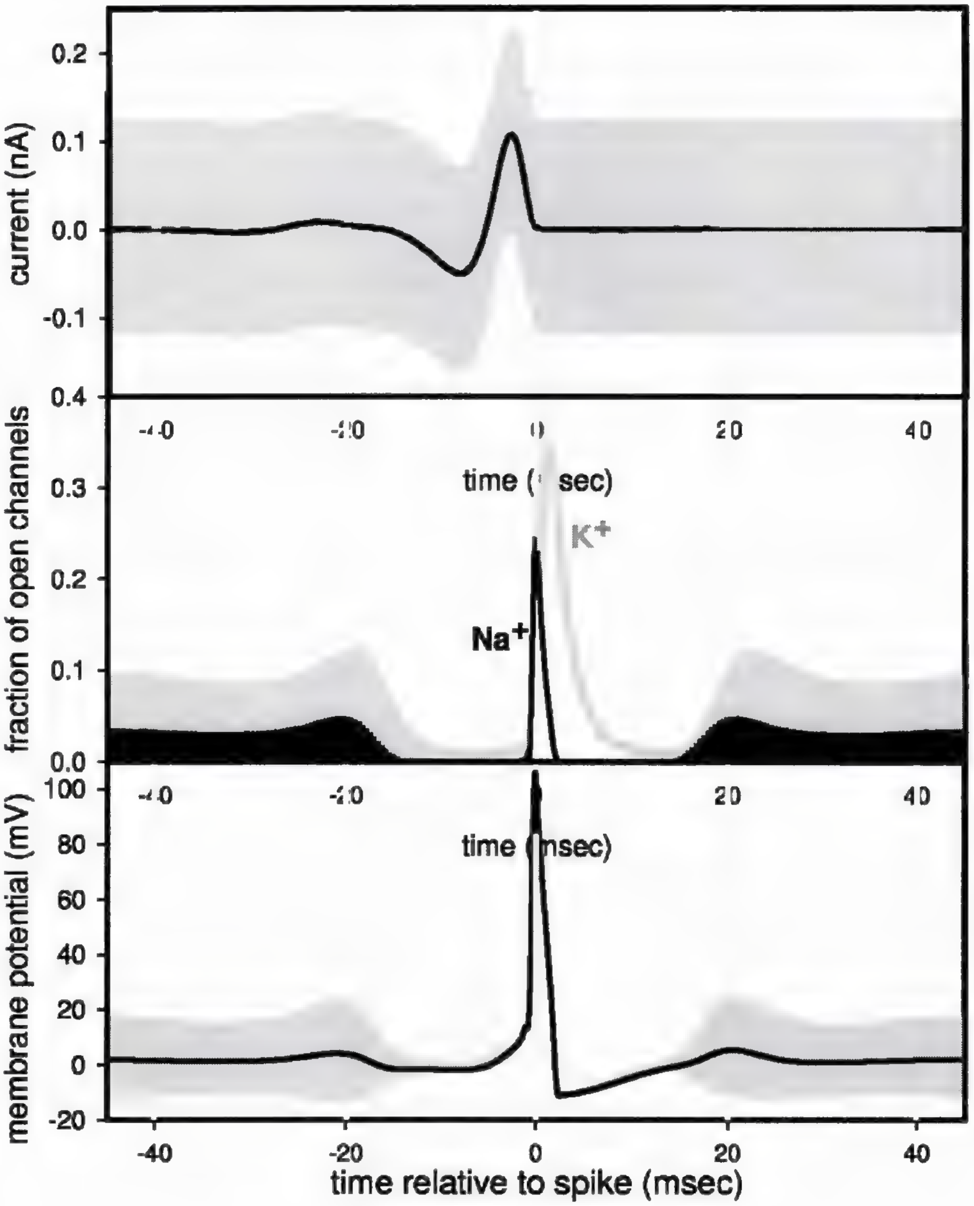


Figure 1: Spike-triggered averages with standard deviations for (top) the input current I, (middle) the fraction of open K^+ and Na^+ channels, and (bottom) the membrane voltage V, for the parameter regime $I_0 = 0$ and $S = 6.50 \times 10^{-4}$ nA^2 sec.

other hand, there is no obvious bottleneck in the current trajectories, so that the current variance is almost constant throughout the spike. This is qualitatively consistent with the idea of dimensionality reduction: if the neuron ignores most of the dimensions along which the current can vary, then the variance, which is shared almost equally among all dimensions for this near white noise, can change by only a small amount.

4.2 Interspike Interaction. Although the STA has the form of a differentiating kernel, suggesting that the neuron detects edge-like events in the current versus time, there must be a DC component to the cell's response. We recall that for constant inputs, the HH model undergoes a bifurcation to constant frequency spiking, where the frequency is a function of the value of the input above onset. Correspondingly, the STA does not sum precisely to zero; one might think of it as having a small integrating component that allows the system to spike under DC stimulation, albeit only above a threshold.

The system's tendency to periodic spiking under DC current input also is felt under dynamic stimulus conditions and can be thought of as a strong interaction between successive spikes. We illustrate this by considering a different parameter regime with a small DC current and some added noise ($I_0 = 0.11$ nA and $S = 0.8 \times 10^{-4}$ nA^2 sec). Note that the DC component puts the neuron in the metastable region of its $f - I$ curve (see Figure 2). In this regime, the neuron tends to fire quasi-regular trains of spikes intermittently, as shown in Figure 3. We will refer to these quasi-regular spike sequences as "bursts" (note that this term is often used to refer to compound spikes in neurons with additional channels; such events do not occur in the HH model).

Spikes can be classified into three types: those initiating a spike burst, those within a burst, and those ending a burst. The minimum length of

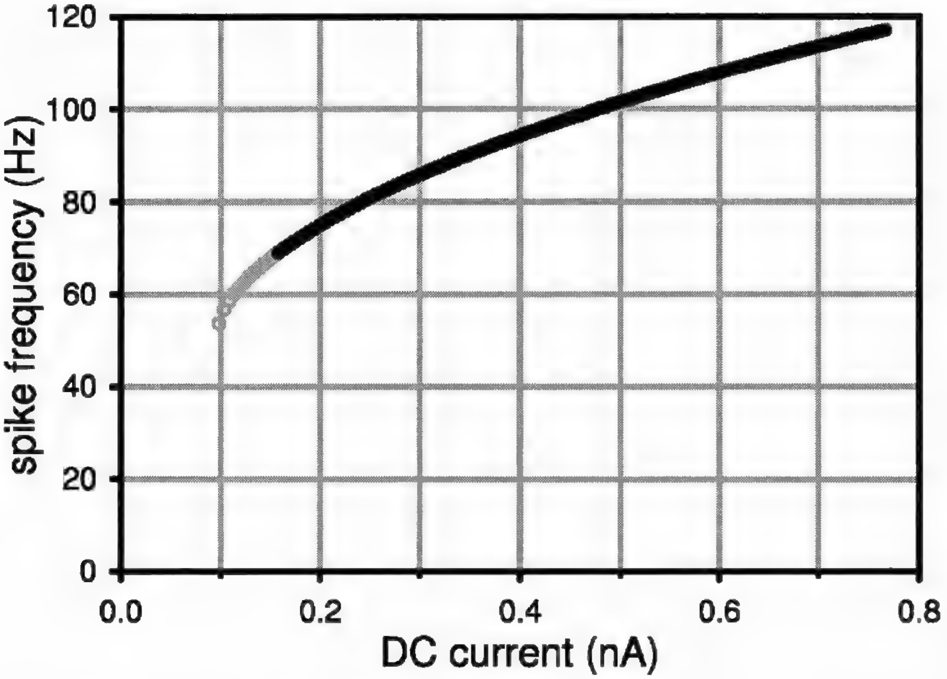


Figure 2: Firing rate of the HH neuron as a function of injected DC current. The empty circles at moderate currents denote the metastable region, where the neuron may be either spiking or silent.

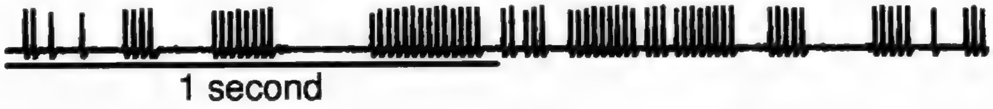


Figure 3: Segment of a typical spike train in a "bursting" regime.

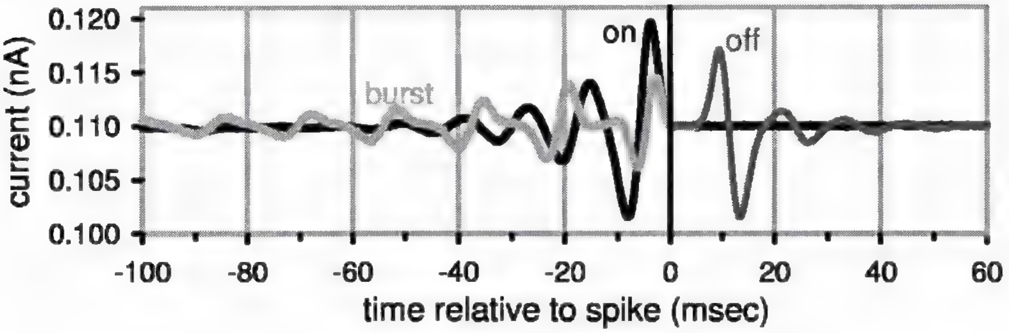


Figure 4: Spike-triggered averages, derived from spikes leading ("on"), inside ("burst"), and ending ("off") a burst. The parameters of this bursting regime are $I_0 = 0.11$ nA and $S = 0.8 \times 10^{-4}$ nA^2 sec. Note that the burst-ending spike average is, by construction, identical to that of any other within-burst spike for $t < 0$.

the silence between bursts is taken in this case to be 70 msec. Taking these three categories of spike as different "symbols" (de Ruyter van Steveninck & Bialek, 1988), we can determine the average stimulus for each. These are shown in Figure 4 with the spike at $t = 0$.

In this regime, the initial spike of a burst is preceded by a rapid oscillation in the current. Spikes within a burst are affected much less by the current; the feature immediately preceding such spikes is similar in shape to a single "wavelength" of the leading spike feature, but is of much smaller amplitude and is temporally compressed into the interspike interval. Hence, although it is clear that the timing of a spike within a burst is determined largely by the timing of the previous spike, the current plays some role in affecting the precise placement. This also demonstrates that the shape of the STA is not the same for all spikes; it depends strongly and nontrivially on the time to the previous spike, and this is related to the observation that subtly different patterns of two or three spikes correspond to very different average stimuli (de Ruyter van Steveninck & Bialek, 1988). For a reader of the spike code, a spike within a burst conveys a different message about the input than the spike at the onset of the burst. Finally, the feature ending a burst has a very similar form to the onset feature, but reversed in time. Thus, to a good approximation, the absence of a spike at the end of a burst can be read as the opposite of the onset of the burst.

In summary, this regime of the HH neuron is similar to a "flip-flop," or 1-bit memory. Like its electronic analog, the neuron's memory is preserved

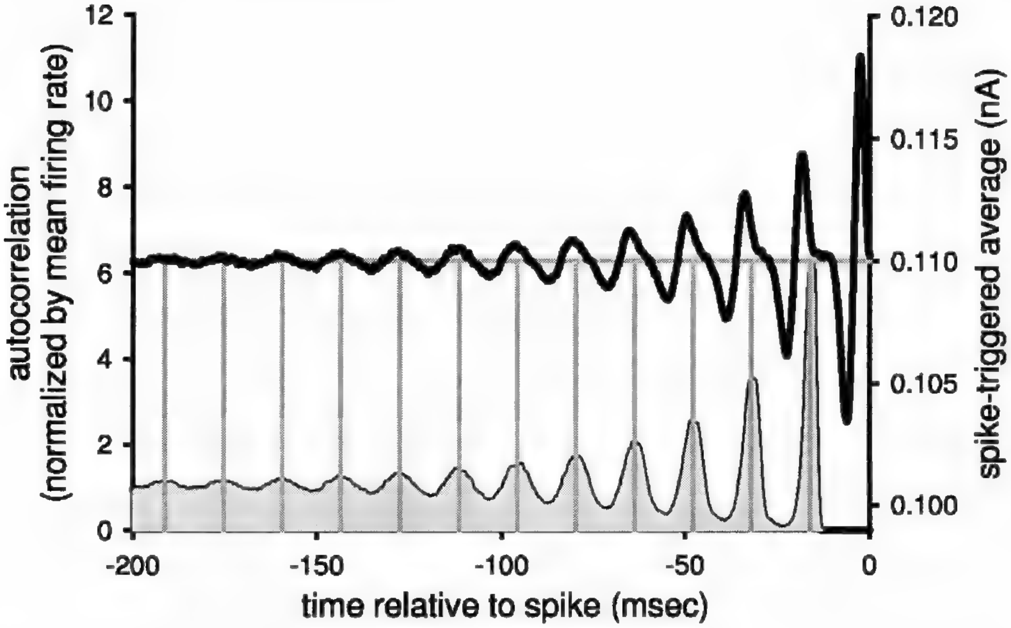


Figure 5: Overall spike-triggered average in the bursty regime, showing the ringing due to the tendency to periodic firing. Plotted in gray is the spike autocorrelation, showing the same oscillations.

by a feedback loop, here implemented by the interspike interaction. Large fluctuations in the input current at a certain frequency "flip" or "flop" the neuron between its silent and spiking states. However, while the neuron is spiking, further details of the input signal are transmitted by precise spike timing within a burst. If we calculate the spike-triggered average of all spikes for this regime, without regard to their position within a burst, then as shown in Figure 5, the relatively well-localized leading spike oscillation of Figure 4 is replaced by a long-lived oscillating function resulting from the spike periodicity. This is shown explicitly by comparing the overall STA with the spike autocorrelation, also shown in Figure 5. This same effect is seen in the STA of the burst spikes, which in fact dominates the overall average. Prediction of spike timing using such an STA would be computationally difficult due to its extension in time, but, more seriously, unsuccessful, as most of the function is an artifact of the spike history rather than the effect of the stimulus.

While the effects of spike interaction are interesting and should be included in a complete model for spike generation, we wish here to consider only the current's role in initiating spikes. Therefore, as we have argued elsewhere, we limit ourselves initially to the cases in which interspike interaction plays no role (Agüera y Arcas et al., 2001; Agüera & Fairhall, 2003). These "isolated" spikes can be defined as spikes preceded by a silent period t_{silence} long enough to ensure decoupling from the timing of the previous spike. A reasonable choice for t_{silence} can be inferred directly from the inter-

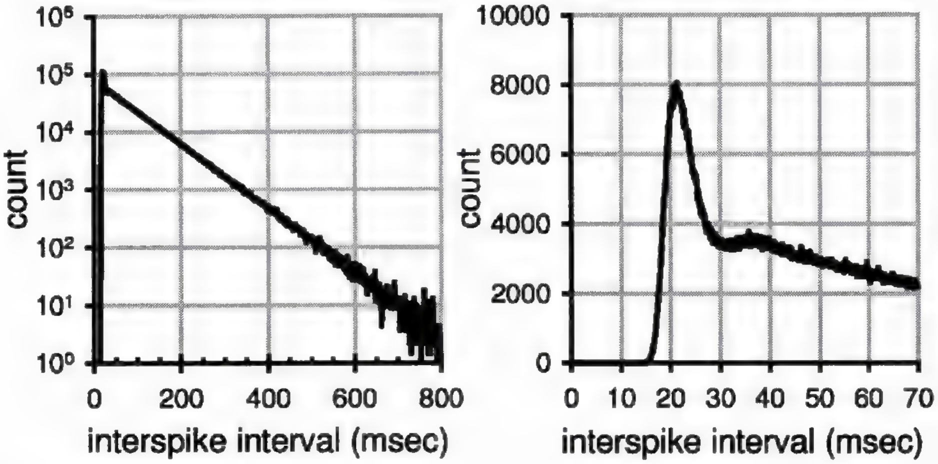


Figure 6: Hodgkin-Huxley interspike interval histogram for the parameters $I_0 = 0$ and $S = 6.5 \times 10^{-4}$ nA^2 sec, showing a peak at a preferred firing frequency and the long Poisson tail. The total number of spikes is $N = 5.18 \times 10^6$. The plot to the right is a closeup in linear scale.

spike interval distribution $P(\Delta t)$, illustrated in Figure 6. For the HH model, as in simpler models and many real neurons (Brenner, Agam, Bialek, & de Ruyter van Steveninck, 1998), the form of $P(\Delta t)$ has three noteworthy features: a refractory "hole" during which another spike is unlikely to occur, a strong mode at the preferred firing frequency, and an exponentially decaying, or Poisson, tail. The details of all three of these features are functions of the parameters of the stimulus (Tiesinga, José, & Sejnowski, 2000), and certain regimes may be dominated by only one or two features. The emergence of Poisson statistics in the tail of the distribution implies that these events are independent, so we can infer that the system has lost memory of the previous spike. We will therefore take isolated spikes to be those preceded by a silent interval $\Delta t \geq t_{\text{silence}}$, where t_{silence} is well into the Poisson regime. The burst-onset spikes of Figure 4 are isolated spikes by this definition.

Note that the bursty behavior evident in Figure 3 is characteristic of "type II" neurons, which begin firing at a well-defined frequency under DC stimulus; "type I" neurons, by contrast, can fire at arbitrarily low frequency under constant input. Nonetheless the analysis that follows is equally applicable to either type of neuron: spikes still interact at close range and become independent at sufficient separation. In what follows, we will be probing the HH neuron with current noise that remains below the DC threshold for periodic firing.

5 Isolated Spike Analysis

Focusing now on isolated spikes, we proceed to a second-order analysis of the current fluctuations around the isolated spike triggered average (see

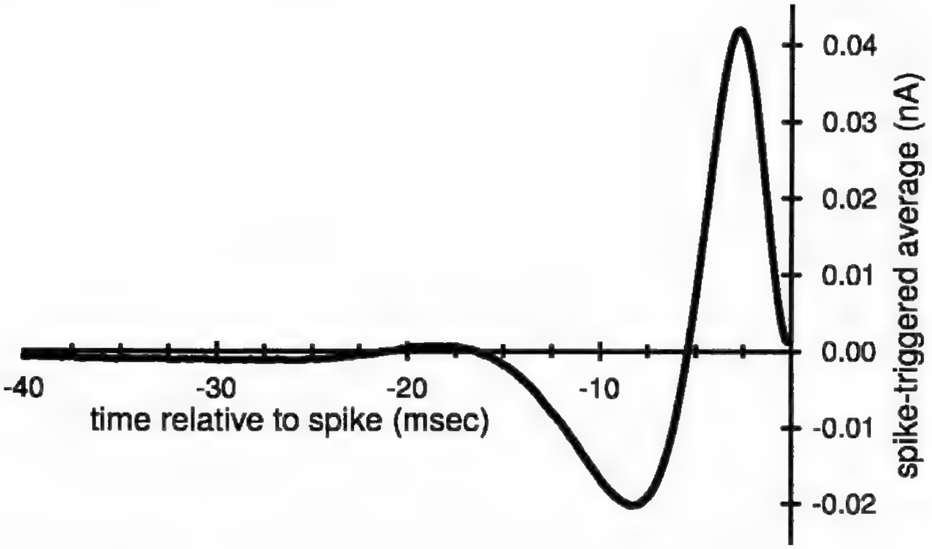


Figure 7: Spike-triggered average stimulus for isolated spikes.

Figure 7). We consider the response of the HH neuron to currents $I(t)$ with mean $I_0 = 0$ and spectral density of $S = 6.5 \times 10^{-4}$ nA^2 sec. Isolated spikes in this regime are defined by $t_{\text{silence}} = 60$ msec.

5.1 How Many Dimensions? As explained in section 2, our path to dimensionality reduction begins with the computation of covariance matrices for stimulus fluctuations surrounding a spike. The matrices are accumulated from stimulus segments 200 samples in length, roughly corresponding to sampling at the timescale sufficiently long to capture the relevant features. Thus, we begin in a 200-dimensional space. We emphasize that the theorem that connects eigenvalues of the matrix ΔC to the number of relevant dimensions is valid only for truly gaussian distributions of inputs and that by focusing on isolated spikes, we are essentially creating a nongaussian stimulus ensemble—namely, those stimuli that generate the silence out of which the isolated spike can appear. Thus, we expect that the covariance matrix approach will give us a heuristic guide to our search for lower-dimensional descriptions, but we should proceed with caution.

The "raw" isolated spike-triggered covariance $C_{\text{iso spike}}$ and the corresponding covariance difference ΔC, equation 2.7, are shown in Figure 8. The matrix shows the effect of the silence as an approximately translationally invariant band preceding the spike, the second-order analog of the constant negative bias in the isolated spike STA (see Figure 7). The spike itself is associated with features localized to ±15 msec. In Figure 9, we show the spectrum of eigenvalues of ΔC computed using a sample of 80,000 spikes. Before calculating the spectrum, we multiply ΔC by C_{prior}^{-1}. This has the effect of giving us eigenvalues scaled in units of the input standard deviation

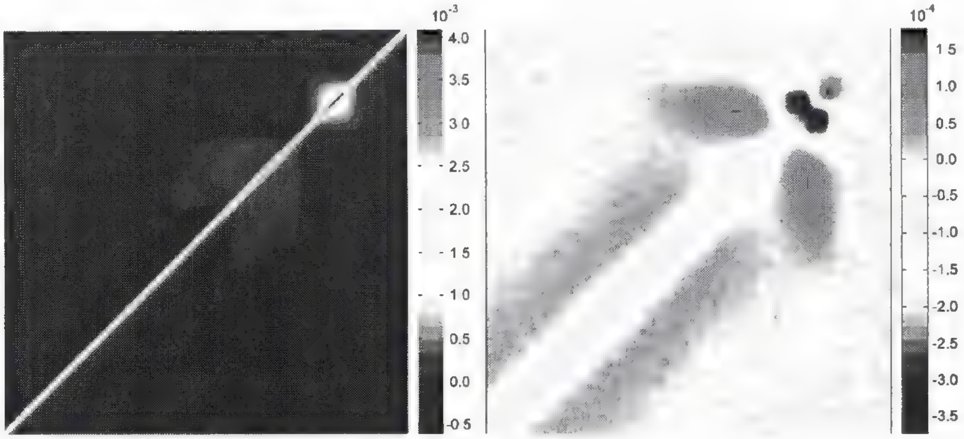


Figure 8: The isolated spike-triggered covariance C_{iso} (left) and covariance difference ΔC (right) for times $-30 < t < 5$ msec. The plots are in units of nA^2.

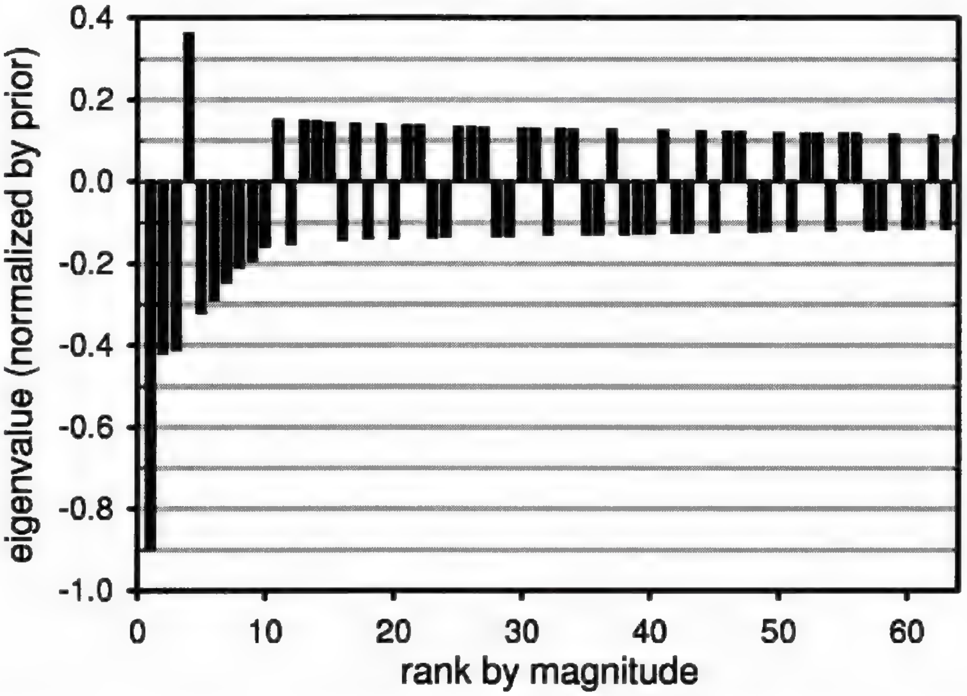


Figure 9: The leading 64 eigenvalues of the isolated spike-triggered covariance after accumulating 80,000 spikes.

along each dimension. Because the correlation time is short, C_{prior} is nearly diagonal.

While the eigenvalues decay rapidly, there is no obvious set of outstanding eigenvalues. To verify that this is not an effect of finite sampling, Figure 10 shows the spectrum of eigenvalue magnitudes as a function of sample size N. Eigenvalues that are truly zero up to the noise floor determined by

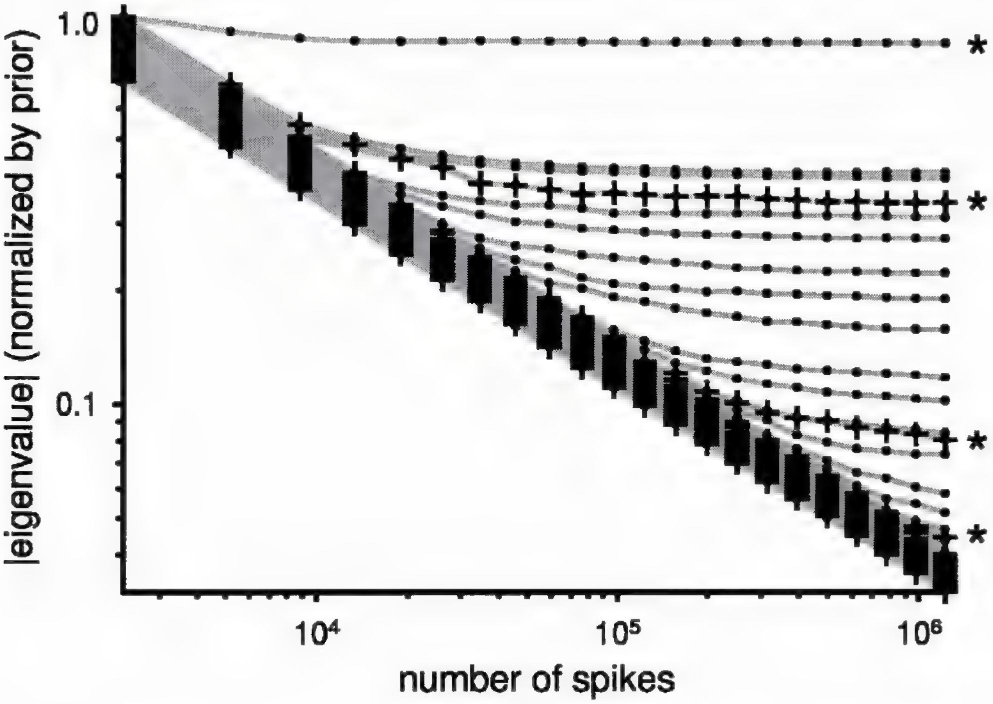


Figure 10: Convergence of the leading 64 eigenvalues of the isolated spike-triggered covariance with increasing sample size. The log slope of the diagonal is $1/\sqrt{n_{\text{spikes}}}$. Positive eigenvalues are indicated by crosses and negative by dots. The spike-associated modes are labeled with an asterisk.

sampling decrease like $\sqrt{N}$. We find that a sequence of eigenvalues emerges stably from the noise.

These results do not, however, imply that a low-dimensional approximation cannot be identified. The extended structure in the covariance matrix induced by the silence requirement is responsible for the apparent high dimensionality. In fact, as has been shown in Agüera y Arcas and Fairhall (2003), the covariance eigensystem includes modes that are local and spike associated, and others that are extended and silence associated, and thus irrelevant to a causal model of spike timing prediction. Fortunately, because extended silences and spikes are (by definition) statistically independent, there is no mixing between the two types of modes. To identify the spike-associated modes, we follow the diagnostic of Agüera y Arcas and Fairhall (2003), computing the fraction of the energy of each mode concentrated in the period of silence, which we take to be $-60 \leq t \leq -40$ msec. The energy of a spike-associated mode in the silent period is due entirely to noise and will therefore decrease like $1/n_{\text{spikes}}$ with increasing sample size, while this energy remains of order unity for silence modes. Carrying out the test on the covariance modes, we obtain Figure 11, which shows that the first and fourth modes rapidly emerge as spike associated. Two further spike-associated modes appear over the sample shown, with the suggestion

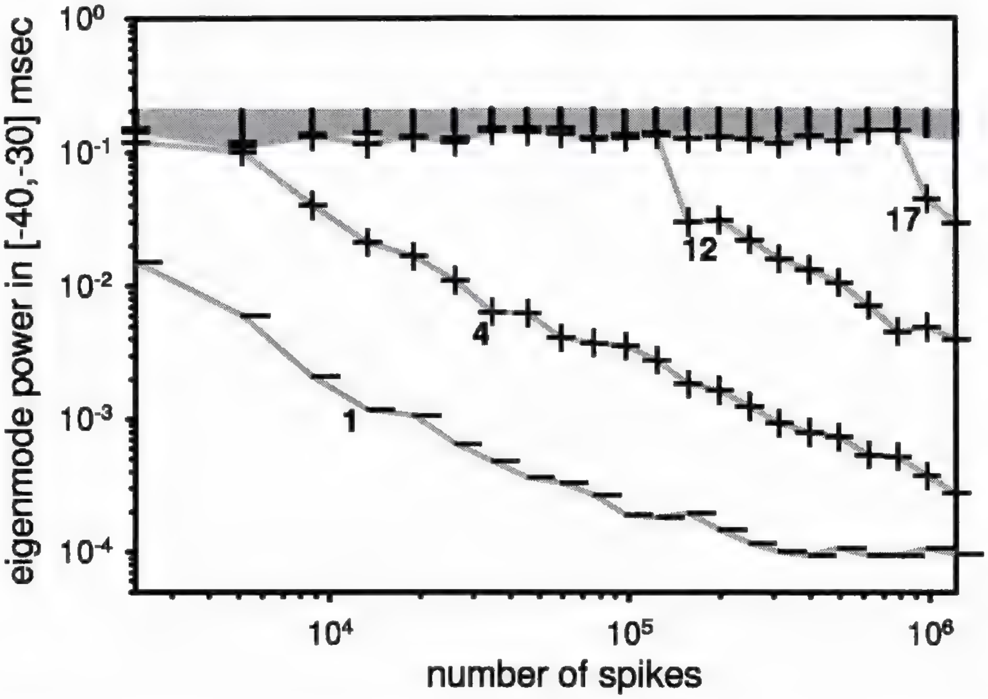


Figure 11: For the leading 64 modes, fraction of the mode energy over the interval $-40 < t < -30$ msec as a function of increasing sample size. Modes emerging with low energy are spike associated. The symbols indicate the sign of the eigenvalue.

of other, weaker modes yet to emerge. The two leading silence modes are shown in Figure 12. Those shown are typical; most modes resemble Fourier modes, as the silence condition is close to time translationally invariant.

Examining the eigenvectors corresponding to the two leading spike-associated eigenvalues, which for convenience we will denote s_1 and s_2 (although they are not the leading modes of the matrix), we find (see Figure 13) that the first mode closely resembles the isolated spike STA and the second is close to the derivative of the first. Both modes approximate differentiating operators; there is no linear combination of these modes that would produce an integrator.

If the neuron filtered its input and generated a spike when the output of the filter crosses threshold, we would find two significant dimensions associated with a spike. The first dimension would correspond simply to the filter, as the variance in this dimension is reduced to zero (for a noiseless system) at the occurrence of a spike. As the threshold is always crossed from below, the stimulus projection onto the filter's derivative must be positive, again resulting in a reduced variance. It is tempting to suggest, then, that filtered threshold crossing is a good approximation to the HH model, but we will see that this is not correct.

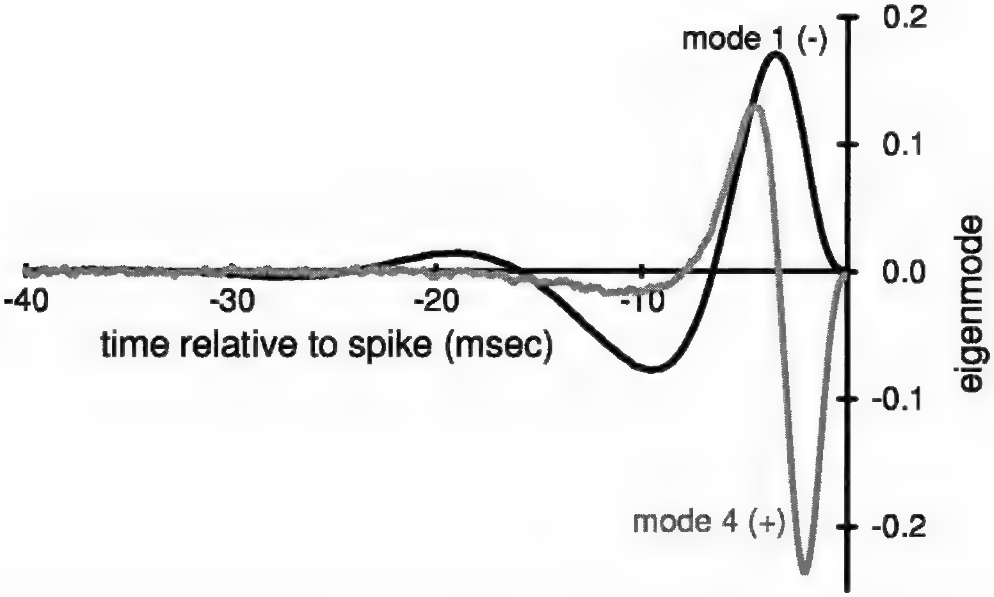


Figure 12: Modes 2 and 3 of the spike-triggered covariance (silence associated).

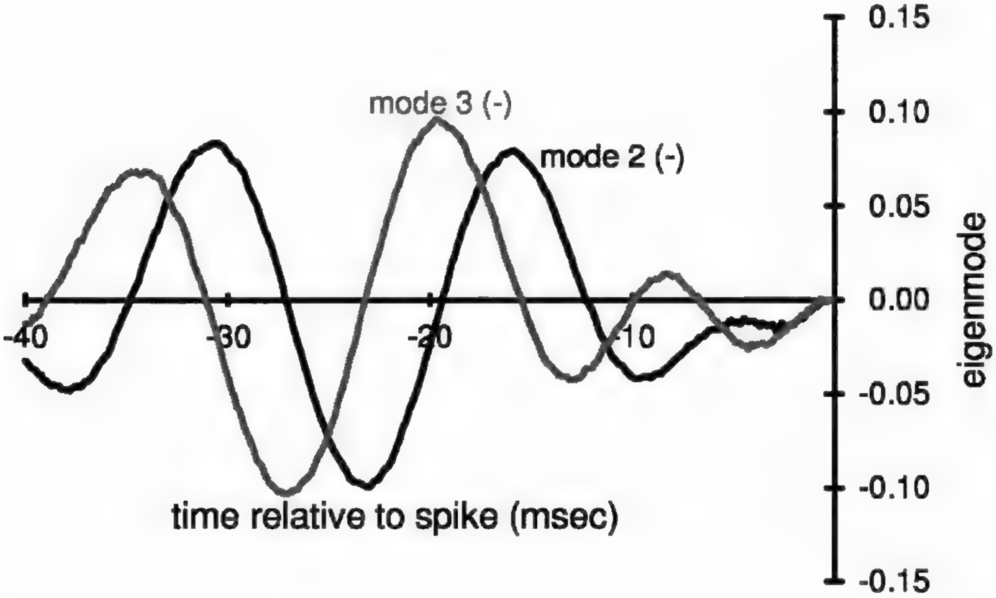


Figure 13: Modes 1 and 4 of the spike-triggered covariance, which are the leading spike-associated modes.

5.2 Evaluating the Nonlinearity. At each instant of time, we can find the projections of the stimulus along the leading spike-associated dimensions s_1 and s_2. By construction, the distribution of these signals over the whole experiment, $P(s_1, s_2)$, is gaussian. The appropriate prior for the isolation condition, $P(s_1, s_2 \mid \text{silence})$, differs only subtly from the gaussian prior. On the other hand, for each spike, we obtain a sample from the dis-

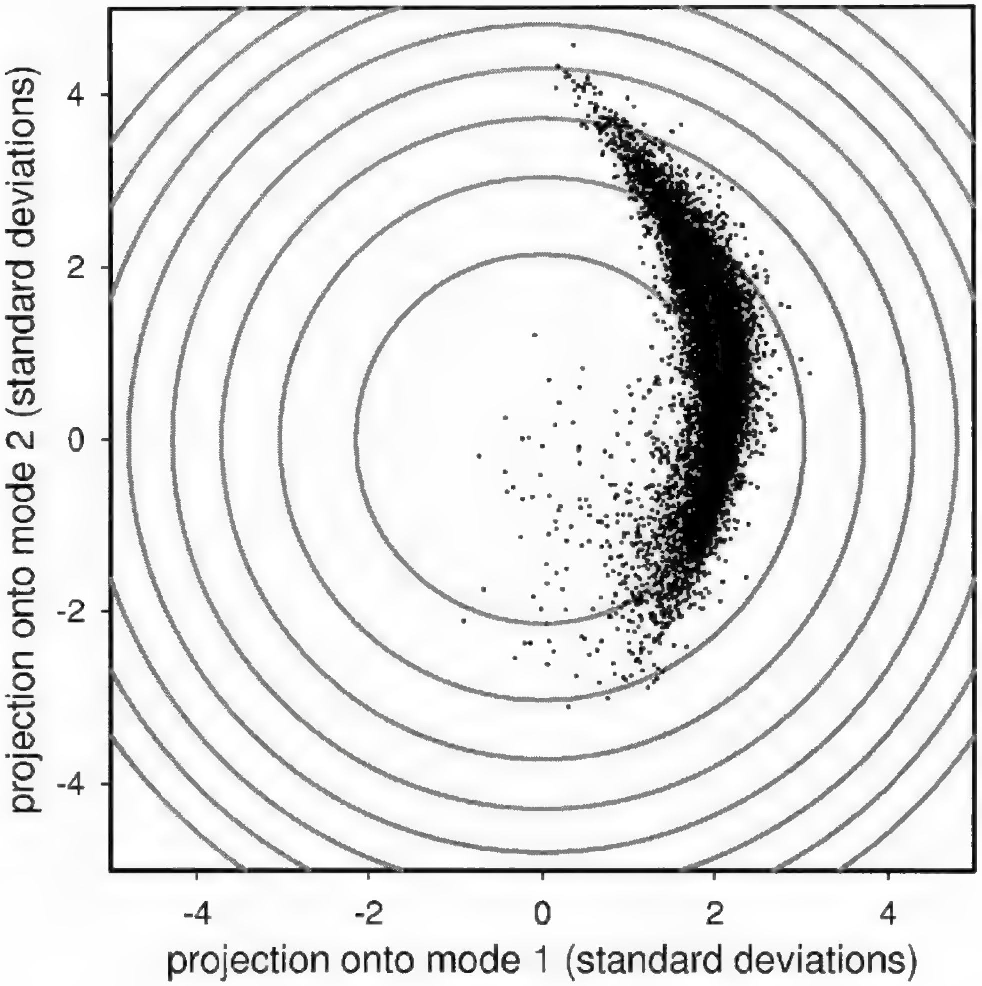


Figure 14: 10^4 spike-conditional stimuli (or "spike histories") projected along the first two covariance modes. The axes are in units of standard deviation on the prior gaussian distribution. The circles, from the inside out, enclose all but $10^{-1}, 10^{-2}, \ldots, 10^{-8}$ of the prior.

tribution $P(s_1, s_2 \mid \text{iso spike at } t_0)$, leading to the picture in Figure 14. The prior and spike-conditional distributions are clearly better separated in two dimensions than in one, which means that the two-dimensional description captures more information than projection onto the spike-triggered average alone. Surprisingly, the spike-conditional distribution is curved, unlike what we would expect for a simple thresholding device. Furthermore, the eigenvalue of ΔC, which we associate with the direction of threshold crossing (plotted on the y-axis in Figure 14), is positive, indicating increased rather than decreased variance in this direction. As we see, projections onto this mode are almost equally likely to be positive or negative, ruling out the threshold crossing interpretation.

Combining equations 2.2 and 2.3 for isolated spikes, we have

$$g(s_1, s_2) = \frac{P(s_1, s_2 \mid \text{iso spike at } t_0)}{P(s_1, s_2 \mid \text{silence})}, \tag{5.1}$$

so that these two distributions determine the input-output relation of the neuron in this 2D space (Brenner, Bialek, et al., 2000). Recall that although the subspace is linear, g can have arbitrary nonlinearity. Figure 14 shows that this input-output relation has clear structure, but also some fuzziness. As the HH model is deterministic, the input-output relation should be a singular function in the continuous space of inputs—spikes occur only when certain exact conditions are met. Of course, finite time resolution introduces some blurring, and so we need to understand whether the blurring of the input-output relation in Figure 14 is an effect of finite time resolution or a real limitation of the 2D description.

5.3 Information Captured in Two Dimensions. We will measure the effectiveness of our description by computing the information in the 2D approximation, according to the methods described in section 3. If the two-dimensional approximation were exact, we would find that $I^{s_1,s_2}_{\text{iso spike}} = I_{\text{iso spike}}$; more generally, one finds $I^{s_1,s_2}_{\text{iso spike}} \leq I_{\text{iso spike}}$, and the fraction of the information captured measures the quality of the approximation. This fraction is plotted in Figure 15 as a function of time resolution. For comparison, we also show the information captured in the one-dimensional case, considering only the stimulus projection along the STA.

We find that our low-dimensional model captures a substantial fraction of the total information available in spike timing in an HH neuron over a range of time resolutions. The approximation is best near $\Delta t = 3$ msec,

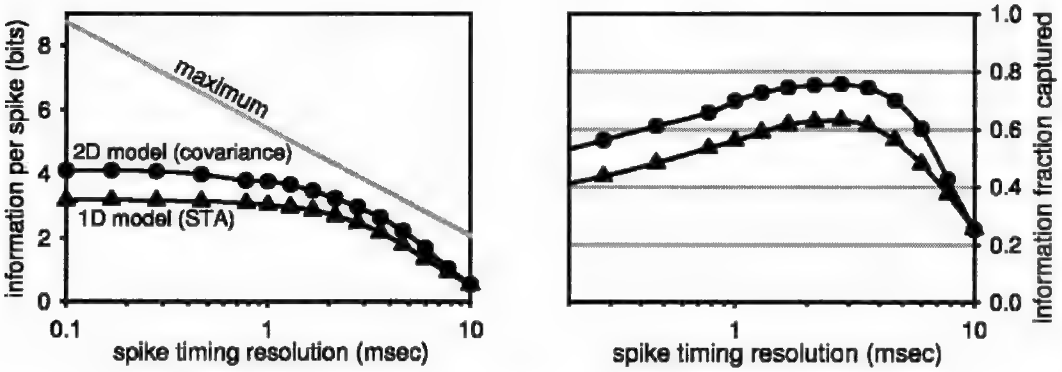


Figure 15: Bits per spike (left) and fraction of the theoretical limit (right) of timing information in a single spike at a given temporal resolution captured by projection onto the STA alone (triangles) and projection onto ΔC covariance modes 1 and 2 (circles).

reaching 75%. Thus, the complex nonlinear dynamics of the HH model can be approximated by saying that the neuron is sensitive to a 2D linear subspace in the high-dimensional space of input signals, and this approximate description captures up to 75% of the mutual information between input currents and (isolated) spike arrival times.

The dependence of information on time resolution (see Figure 15) shows that the absolute information captured saturates for both the 1D and 2D cases, at ≈ 3.2 and 4.1 bits respectively. Hence, for smaller Δt, the information fraction captured drops. The model provides, at its best, a time resolution of 3 msec, so that information carried by more precise spike timing is lost in our low-dimensional projection. Might this missing information be important for a real neuron? Stochastic HH simulations with realistic channel densities suggest that the timing of spikes in response to white noise stimuli is reproducible to within 1 to 2 msec (Schneidman, Freedman, & Segev, 1998), a figure that is comparable to what is observed for pyramidal cells in vitro (Mainen & Sejnowski, 1995), as well in vivo in the fly's visual system (de Ruyter van Steveninck, Lewen, Strong, Koberle, & Bialek, 1997; Lewen, Bialek, & de Ruyter van Steveninck, 2001), the vertebrate retina (Berry, Warland, & Meister, 1997), the cat lateral geniculate nucleus (LGN) (Reinagel & Reid, 2000), and the bat auditory cortex (Dear, Simmons, & Fritz, 1993). This suggests that such timing details may indeed be important. We must therefore ask why our approximation seems to carry an inherent time resolution limitation and why, even at its optimal resolution, the full information in the spike is not recovered.

For many purposes, recovering 75% of the information at ~ 3 msec resolution might be considered a resounding success. On the other hand, with such a simple underlying model, we would hope for a more compelling conclusion. From a methodological point of view, it behooves us to ask what we are missing in our 2D model, and perhaps the methods we use in finding the missing information in the present case will prove applicable more generally.

6 What Is Missing?

The obvious first approach to improving the 2D approximation is to add more dimensions. Let us consider the neglected modes. We recall from Figure 11 that in simulations with very large numbers of spikes, we can isolate at least two more modes that have significant eigenvalues and are associated with the isolated spike rather than the preceding silence; these are shown in Figure 16. We see that these modes look like higher-order derivatives, which makes some sense since we are missing information at high time resolution. On the other hand, if all we are doing is expanding in a basis of higher-order derivatives it is not clear that we will do qualitatively better by including one or two more terms, particularly given that sampling higher-dimensional distributions becomes very difficult.

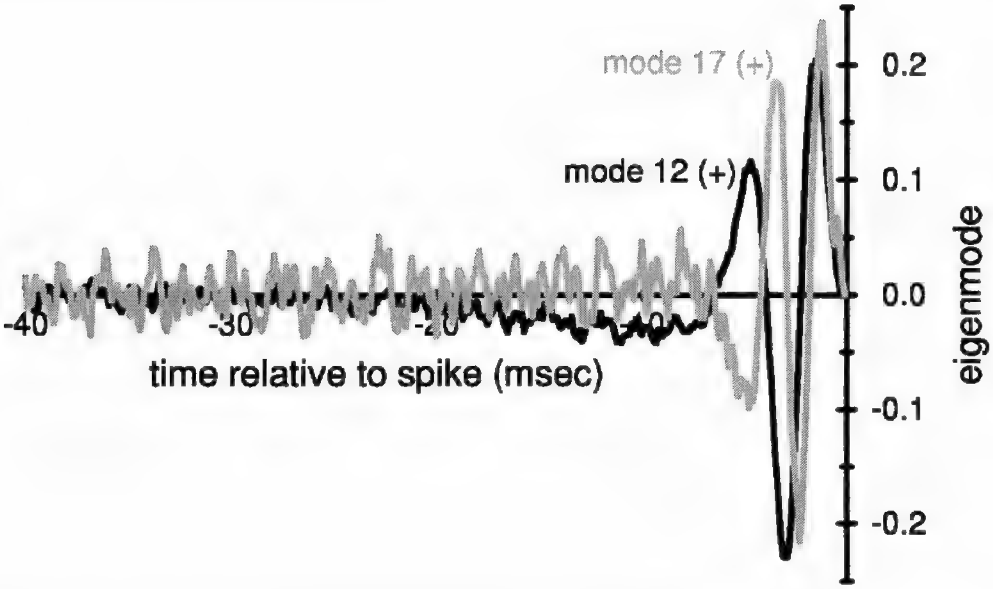


Figure 16: The two next spike-associated modes. These resemble higher-order derivatives.

Our original model attempted to approximate the set of relevant features as lying within a K-dimensional linear subspace of the original D-dimensional input space. The covariance spectrum indicates that additional dimensions play a small but significant role. We suggest that a reasonable next step is to consider the relevant feature space to be low dimensional but not flat. The first two covariance modes define a plane; we will consider next a 2D geometric construction that curves into additional dimensions.

Several methods have been proposed for the general problem of identifying low-dimensional nonlinear manifolds (Cottrell, Munro, & Zipser, 1988; Boser, Guyon, & Vapnik, 1992; Guyon, Boser, & Vapnik, 1993; Oja & Karhunen, 1995; Roweis & Saul, 2000), but these various approaches share the disadvantage that the manifold or, equivalently, the relevant set of features remains implicit. Our hope is to understand the behavior of the neuron explicitly; we therefore wish to obtain an explicit representation of this curved feature space in terms of the basis vectors (features) that span it.

A first approach to determining the curved subspace is to approximate it as a set of locally linear "tiles." At any place on the surface, we wish to find two orthogonal directions that form the surface's local linear approximation. Curvature of the subspace means that these two dimensions will vary across the surface. We apply a simple algorithm to construct a locally linear tiling with the advantage that it requires only first-order statistics. First, we take one of the directions to be globally defined; it is natural to take this direction to be the overall spike-triggered average. Allowing for curvature in the feature subspace means that along the direction of the STA, we allow the second, orthogonal direction to vary. In principle, this vari-

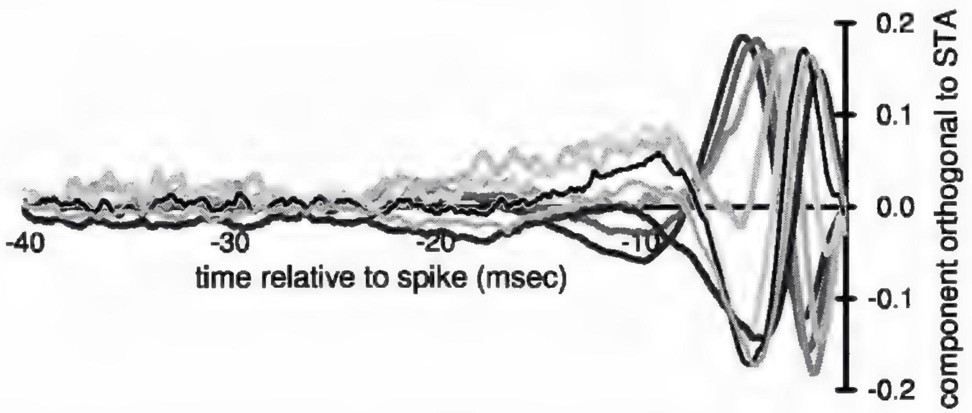


Figure 17: The orthonormal components of spike-triggered averages from 80,000 spikes conditioned on their projection onto the overall spike-triggered average (eight conditional averages shown).

ation is continuous, but we will not be able to sample sufficiently to find the complete description of the second direction, so we sort the stimulus histories according to their projection along this direction and bin the sorted histories into a small number of bins. This defines the number of tiles used to cover the surface. We determine the conditional average of the stimulus histories in each bin and compute the (normalized) component orthogonal to the overall STA. This provides a second locally meaningful basis vector for the subspace in that bin. The resulting family of curves orthogonal to the STA is shown in Figure 17. Applying singular value decomposition to the family of curves shows that there are at least four significant independent directions in stimulus space apart from the STA. This gives us an estimate of the embedding dimension of the feature subspace.

Geometrically, this construction is equivalent to approximating the surface as a twisting ribbon, with the leading direction that of the STA, but where the surface is allowed to rotate about the STA axis. Further, we have discretized the twisting direction into a small number of tiles. The discretization is fixed by the size of the data. There is a trade-off between the precision of estimating the conditional average and the fidelity with which one follows the twist. Here we have restricted ourselves to an experimentally realistic number of isolated spikes (80,000), using an equal number of spikes per bin. Varying over the number of bins, we found that eight bins gave the best result. Note that our model is discontinuous; it might be possible to improve it by interpolating smoothly between successive tiles.

Computing the information as a function of Δt using this locally linear model, we obtain the curve shown in Figure 18, where the results can be compared against the information found from the STA alone and from the covariance modes. The information from the new model captures a maxi-

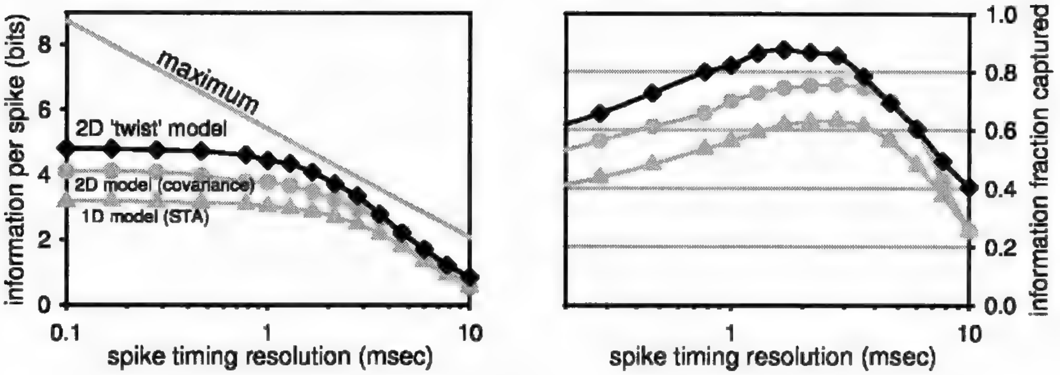


Figure 18: Bits per spike (left) and fraction of the theoretical limit (right) of timing information in a single spike at a given temporal resolution captured by the locally linear tiling "twist" model (diamonds), compared to models using the STA alone (triangles), and projection onto ΔC covariance modes 1 and 2 (circles).

mum of 4.8 bits, recovering $\sim 90\%$ of the information at a time resolution of approximately 1 msec.

One of the main strengths of this simple approach is that we have succeeded in extracting additional geometrical information about the feature subspace using very limited data, as we compute only averages. Note that a similar number of spikes cannot resolve more than two spike-associated covariance modes in the covariance matrix analysis.

7 Discussion

The HH equations describe the dynamics of four degrees of freedom, and almost since these equations were first written down, there have been attempts to find simplifications or reductions. FitzHugh and Nagumo proposed a 2D system of equations that approximate the HH model (Fitzhugh, 1961; Nagumo, Arimoto, & Yoshikawa, 1962), and this has the advantage that one can visualize the trajectories directly in the plane and thus achieve an intuitive graphical understanding of the dynamics and its dependence on parameters. The need for reduction in the sense pioneered by FitzHugh and by Nagumo et al. has become only more urgent with the growing use of increasingly complex HH-style model neurons with many different channel types. With this problem in mind, Kepler, Abbott, and Marder have introduced reduction methods that are more systematic, making use of the difference in timescales among the gating variables (Kepler, Abbott, & Marder, 1992; Abbott & Kepler, 1990).

In the presence of constant current inputs, it makes sense to describe the HH equations as a 4D autonomous dynamical system; by well-known methods in dynamical systems theory, one could consider periodic input currents by adding an extra dimension. The question asked by FitzHugh and Nagumo was whether this 4D or 5D description could be reduced to two or three dimensions.

Closer in spirit to our approach is the work by Kistler, Gerstner, and van Hemmen (1997), who focused in particular on the interaction among successive action potentials. They argued that one could approximate the HH model by a nearly linear dynamical system with a threshold, identifying threshold crossing with spiking, provided that each spike generated either a change in threshold or an effective input current that influences the generation of subsequent spikes.

The notion of model dimensionality considered here is distinct from the dynamical systems perspective in which one simply counts the system's degrees of freedom. Here we are attempting to find a description of the dynamics, which is essentially functional or computational. We have identified the output of the system as spike times, and our aim is to construct as complete a description as possible of the mapping between input and output. The dimensionality of our model is that of the space of inputs relevant for this mapping. There is no necessary relationship between these two notions of dimensionality. For example, in a neural network with two attractors, a system described by a potentially large number of variables, there might be a simple rule (perhaps even a linear filter) that allows us to look at the inputs to the network and determine the times at which the switching events will occur. Conversely, once we leave the simplified world of constant or periodic inputs, even the small number of differential equations describing a neuron's channel dynamics could in principle be equivalent to a very complicated set of rules for mapping inputs into spike times.

In our context, simplicity is (roughly) feature selectivity: the mapping is simple if spiking is determined by a small number of features in the complex history of inputs. Following the ideas that emerged in the analysis of motion-sensitive neurons in the fly (de Ruyter van Steveninck & Bialek, 1988; Bialek & de Ruyter van Steveninck, 2003), we have identified "features" with "dimensions" and searched for low-dimensional descriptions of the input history that preserve the mutual information between inputs and outputs (spike times). We have considered only the generation of isolated spikes, leaving aside the question of how spikes interact with one another, as considered by Kistler et al. (1997). For these isolated spikes, we began by searching for projections onto a low-dimensional linear subspace of the originally $\sim$ 200-dimensional stimulus space, and we found that a substantial fraction of the mutual information could be preserved in a model with just two dimensions. Searching for the information that is missing from this model, we found that rather than adding more (Euclidean) dimensions, we could capture approximately 90% of the information at high time resolution by keeping a 2D description but allowing these dimensions to vary over the surface, so that the neuron is sensitive to stimulus features that lie in a curved 2D subspace.

The geometrical picture of neurons as being sensitive to features that are defined by a low-dimensional stimulus subspace is attractive and, as noted in section 1, corresponds to a widely shared intuition about the nature of neuronal feature selectivity. While curved subspaces often appear as the targets for learning in complex neural computations such as invariant object recognition, the idea that such subspaces appear already in the description of single-neuron computation we believe to be novel.

While we have exploited the fact that long simulations of the HH model are quite tractable to generate large amounts of "data" for our analysis, it is important that, in the end, our construction of a curved, relevant, stimulus subspace involves a series of computations that are just simple generalizations of the conventional reverse correlation or spike-triggered average. This suggests that our approach can be applied to real neurons without requiring qualitatively larger data sets than might have been needed for a careful reverse correlation analysis. In the same spirit, recent work has shown how covariance matrix analysis of the fly's motion-sensitive neurons can reveal nonlinear computations in a 4D subspace using data sets of fewer than 10^4 spikes (Bialek & de Ruyter van Steveninck, 2003). Low-dimensional linear subspaces can be found even in the response of model neurons to naturalistic inputs if one searches directly for dimensions that capture the largest fraction of the mutual information between inputs and spikes (Sharpee et al., in press), and again the errors involved in identifying the relevant dimensions are comparable to the errors in reverse correlation (Sharpee, Rust, & Bialek, 2003). All of these results point to the practical feasibility of describing real neurons in terms of nonlinear computation on low-dimensional relevant subspaces in a high-dimensional stimulus space.

Our reduced model of the HH neuron both illustrates a novel approach to dimensional reduction and gives new insight into the computation performed by the neuron. The reduced model is essentially that of an edge detector for current trajectories but is sensitive to a further stimulus parameter, producing a curved manifold. An interpretation of this curvature will be presented in a forthcoming manuscript. This curved representation is able to capture almost all information that isolated spikes convey about the stimulus, or conversely, allow us to predict isolated spike times with high temporal precision from the stimulus. The emergence of a low-dimensional curved manifold in a model as simple as the HH neuron suggests that such a description may also be appropriate for biological neurons.

Our approach is limited in that we address only isolated spikes. This restricted class of spikes nonetheless has biological relevance; for example, in vertebrate retinal ganglion cells (Berry & Meister, 1999), in rat somatosensory cortex (Panzeri, Petersen, Schultz, Lebedev, & Diamond, 2001), and in LGN (Reinagel, Godwin, Sherman, & Koch, 1999), the first spike of a burst has been shown to convey distinct (and the majority of the) information.

However, a clear next step in this program is to extend our formalism to take into account interspike interaction. For neurons or models with explicit long timescales, adaptation induces very long-range history dependence, which complicates the issue of spike interactions considerably. A full understanding of the interaction between stimulus and spike history will therefore in general involve understanding the meanings of spike patterns (de Ruyter van Steveninck & Bialek, 1988; Brenner, Strong, et al., 2000) and the influence of the larger statistical context (Fairhall et al., 2001). Our results point to the need for a more parsimonious description of self-excitation, even for the simple case of dependence on only the last spike time.

We close by reminding readers of the more ambitious goal of building bridges between the burgeoning molecular-level description of neurons and the functional or computational level. Armed with a description of spike generation as a nonlinear operation on a low-dimensional, curved manifold in the space of inputs, it is natural to ask how the details of this computational picture are related to molecular mechanisms. Are neurons with more different types of ion channels sensitive to more stimulus dimensions, or do they implement more complex nonlinearities in a low-dimensional space? Are adaptation and modulation mechanisms that change the nonlinearity separable from those that change the dimensions to which the cell is sensitive? Finally, while we have shown how a low-dimensional description can be constructed numerically from observations of the input-output properties of the neuron, one would like to understand analytically why such a description emerges and whether it emerges universally from the combinations of channel dynamics selected by real neurons.

Acknowledgments

We thank N. Brenner for discussions at the start of this work, and M. Berry for comments on the manuscript.

References

Abbott, L. F., & Kepler, T. (1990). Model neurons: From Hodgkin-Huxley to Hopfield. In L. Garrido (Ed.), *Statistical mechanisms of neural networks* (pp. 5–18). Berlin: Springer-Verlag.

Agüera y Arcas, B. (1998). *Reducing the neuron: A computational approach.* Unpublished master's thesis, Princeton University.

Agüera y Arcas, B., Bialek, W., & Fairhall, A. L. (2001). What can a single neuron compute? In T. Leen, T. Dietterich, & V. Tresp (Eds.), *Advances in neural information processing systems, 13* (pp. 75–81). Cambridge, MA: MIT Press.

Agüera y Arcas, B., & Fairhall, A. (2003). What causes a neuron to spike? *Neural Computation, 15*, 1789–1807.

Barlow, H. B. (1953). Summation and inhibition in the frog's retina. *J. Physiol., 119*, 69–88.

Barlow, H. B., Hill, R. M., & Levick, W. R. (1964). Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *J. Physiol., 173*, 377–407.

Berry, M. J. II, & Meister, M. (1999). The neural code of the retina. *Neuron, 22*, 435–450.

Berry, M. J. II, Warland, D., & Meister, M. (1997). The structure and precision of retinal spike trains. *Proc. Natl. Acad. Sci. U.S.A., 94*, 5411–5416.

Bialek, W., & de Ruyter van Steveninck, R. R. (2003). *Features and dimensions: Motion estimation in fly vision.* Unpublished manuscript.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In D. Haussler (Ed.), *5th Annual ACM Workshop on COLT* (pp. 144–152). Pittsburgh, PA: ACM Press.

Bray, D. (1995). Protein molecules as computational elements in living cells. *Nature, 376*, 307–312.

Brenner, N., Agam, O., Bialek, W., & de Ruyter van Steveninck, R. R. (1998). Universal statistical behavior of neural spike trains. *Phys. Rev. Lett., 81*, 4000–4003.

Brenner, N., Bialek, W., & de Ruyter van Steveninck, R. R. (2000). Adaptive rescaling maximizes information transmission. *Neuron, 26*, 695–702.

Brenner, N., Strong, S., Koberle, R., Bialek, W., & de Ruyter van Steveninck, R. R. (2000). Synergy in a neural code. *Neural Comp., 12*, 1531–1552. Available on-line: http://xxx.lanl.gov/abs/physics/9902067.

Cottrell, G. W., Munro, P., & Zipser, D. (1988). Image compression by back propagation: A demonstration of extensional programming. In N. Sharkey (Ed.), *Models of cognition: A review of cognitive science* (Vol. 2, pp. 208–240). Norwood, NJ: Ablex.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory.* New York: Wiley.

de Boer, E., & Kuyper, P. (1968). Triggered correlation. *IEEE Trans. Biomed. Eng., 15*, 169–179.

de Ruyter van Steveninck, R. R., & Bialek, W. (1988). Real-time performance of a movement sensitive in the blowfly visual system: Information transfer in short spike sequences. *Proc. Roy. Soc. Lond. B, 234*, 379–414.

de Ruyter van Steveninck, R., Lewen, G. D., Strong, S. P., Koberle, R., & Bialek, W. (1997). Reproducibility and variability in neural spike trains. *Science, 275*, 1805–1808.

Dear, S. P., Simmons, J. A., & Fritz, J. (1993). A possible neuronal basis for representation of acoustic scenes in auditory cortex of the big brown bat. *Nature, 364*, 620–623.

Fairhall, A., Lewen, G., Bialek, W., & de Ruyter van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature, 412*, 787–792.

Fitzhugh, R. (1961). Impulse and physiological states in models of nerve membrane. *Biophysics J., 1*, 445–466.

Guyon, I. M., Boser, B. E., & Vapnik, V. N. (1993). Automatic capacity tuning of very large VC-dimension classifiers. In S. J. Hanson, J. D. Cowan, & C. Giles (Eds.), *Advances in neural information processing systems, 5* (pp. 147–155). San Mateo, CA: Morgan Kaufmann.

Hartline, H. K. (1940). The receptive fields of optic nerve fibres. *Amer. J. Physiol., 130*, 690–699.

Hille, B. (1992). *Ionic channels of excitable membranes*. Sunderland, MA: Sinauer.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol., 463*, 391–407.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.), 160*, 106–154.

Iverson, L., & Zucker, S. W. (1995). Logical/linear operators for image curves. *IEEE Trans. Pattern Analysis and Machine Intelligence, 17*, 982–996.

Keat, J., Reinagel, P., Reid, R. C., & Meister, M. (2001). Predicting every spike: A model for the responses of visual neurons. *Neuron, 30*(3), 803–817.

Kepler, T., Abbott, L. F., & Marder, E. (1992). Reduction of conductance-based neuron models. *Biological Cybernetics, 66*, 381–387.

Kistler, W., Gerstner, W., & van Hemmen, J. L. (1997). Reduction of the Hodgkin-Huxley equations to a single-variable threshold model. *Neural Computation, 9*, 1015–1045.

Koch, C. (1999). *Biophysics of computation: Information processing in single neurons*. New York: Oxford University Press.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol., 16*, 37–68.

Lewen, G. D., Bialek, W., & de Ruyter van Steveninck, R. R. (2001). Neural coding of naturalistic motion stimuli. *Network, 12*, 317–329.

Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science, 268*, 1503–1506.

Nagumo, J., Arimoto, S., & Yoshikawa, Z. (1962). An active pulse transmission line simulating nerve axon. *Proc. IRE, 50*, 2061–2071.

Oja, E., & Karhunen, J. (1995). Signal separation by nonlinear Hebbian learning. In M. Palaniswami, Y. Attikiouzel, R. J. Marks II, D. Fogel, & T. Fukuda (Eds.), *Computational intelligence—a dynamic system perspective* (pp. 83–97). New York: IEEE Press.

Panzeri, S., Petersen, R., Schultz, S., Lebedev, M., & Diamond, M. (2001). The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron, 29*, 769–777.

Reinagel, P., Godwin, D., Sherman, S. M., & Koch, C. (1999). Encoding of visual information by LGN bursts. *J. Neurophys., 81*, 2558–2569.

Reinagel, P., & Reid, R. C. (2000). Temporal coding of visual information in the thalamus. *J. Neuroscience, 20*(14), 5392–5400.

Rieke, F., Warland, D., Bialek, W., & de Ruyter van Steveninck, R. R. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review, 65*, 386–408.

Rosenblatt, F. (1962). *Principles of neurodynamics*. New York: Spartan Books.

Roweis, S., & Saul, L. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science, 290*, 2323–2326.

Schneidman, E., Freedman, B., & Segev, I. (1998). Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Comp., 10*, 1679–1703.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Sys. Tech. Journal, 27*, 379–423, 623–656.

Sharpee, T., Rust, N. C., & Bialek, W. (in press). Maximally informative dimensions: analysing neural responses to natural signals. *Neural Information Processing Systems 2002*. Available on-line: http://xxx.lanl.gov/abs/physics/0208057.

Sharpee, T., Rust, N. C., & Bialek, W. (2003). *Maximally informative dimensions: Analysing neural responses to natural signals*. Unpublished manuscript.

Stanley, G. B., Lei, F. F., & Dan, Y. (1999). Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *J. Neurosci., 19*(18), 8036–8042.

Theunissen, F., Sen, K., & Doupe, A. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J. Neurosci., 20*, 2315–2331.

Tiesinga, P. H. E., José, J., & Sejnowski, T. (2000). Comparison of current-driven and conductance-driven neocortical model neurons with Hodgkin-Huxley voltage-gated channels. *Physical Review E, 62*, 8413–8419.

Tishby, N., Pereira, F., & Bialek, W. (1999). The information bottleneck method. In B. Hajek & R. S. Sreenivas (Eds.), *Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing* (pp. 368–377). Champaign, IL. Available on-line: http://xxx.lanl.gov/abs/physics/0004057.

Received January 9, 2003; accepted January 28, 2003.

Intrinsic Gain Modulation and Adaptive Neural Coding

Sungho Hong[¤]*, **Brian Nils Lundstrom**, **Adrienne L. Fairhall**

Physiology and Biophysics Department, University of Washington, Seattle, Washington, United States of America

Abstract

In many cases, the computation of a neural system can be reduced to a receptive field, or a set of linear filters, and a thresholding function, or gain curve, which determines the firing probability; this is known as a linear/nonlinear model. In some forms of sensory adaptation, these linear filters and gain curve adjust very rapidly to changes in the variance of a randomly varying driving input. An apparently similar but previously unrelated issue is the observation of gain control by background noise in cortical neurons: the slope of the firing rate versus current (*f-I*) curve changes with the variance of background random input. Here, we show a direct correspondence between these two observations by relating variance-dependent changes in the gain of *f-I* curves to characteristics of the changing empirical linear/nonlinear model obtained by sampling. In the case that the underlying system is fixed, we derive relationships relating the change of the gain with respect to both mean and variance with the receptive fields derived from reverse correlation on a white noise stimulus. Using two conductance-based model neurons that display distinct gain modulation properties through a simple change in parameters, we show that coding properties of both these models quantitatively satisfy the predicted relationships. Our results describe how both variance-dependent gain modulation and adaptive neural computation result from intrinsic nonlinearity.

Citation: Hong S, Lundstrom BN, Fairhall AL (2008) Intrinsic Gain Modulation and Adaptive Neural Coding. PLoS Comput Biol 4(7): e1000119. doi:10.1371/journal.pcbi.1000119

Editor: Karl J. Friston, University College London, United Kingdom

Received January 30, 2008; **Accepted** June 9, 2008; **Published** July 18, 2008

Copyright: © 2008 Hong et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by a Burroughs-Wellcome Careers at the Scientific Interface grant, a Sloan Research Fellowship, and a McKnight Scholar Award to ALF. BNL was supported by grant number F30NS055650 from the National Institute of Neurological Disorders and Stroke, the Medical Scientist Training Program at the University of Washington supported by the National Institute of General Sciences, and an Achievement Rewards for College Scientists fellowship.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: shhong@oist.jp

¤ Current address: Computational Neuroscience Unit, Okinawa Institute of Science and Technology, Onna-son, Okinawa, Japan

Introduction

An *f-I* curve, defined as the mean firing rate in response to a stationary mean current input, is one of the simplest ways to characterize how a neuron transforms a stimulus into a spike train output as a function of the magnitude of a single stimulus parameter. Recently, the dependence of *f-I* curves on other input statistics such as the variance has been examined: the slope of the *f-I* curve, or gain, is modulated in diverse ways in response to different intensities of added noise [1–4]. This enables multiplicative control of the neuronal gain by the level of background synaptic activity [1]: changing the level of the background synaptic activity is equivalent to changing the variance of the noisy balanced excitatory and inhibitory input current to the soma, which modulates the gain of the *f-I* curve. It has been demonstrated that such somatic gain modulation, combined with saturation in the dendrites, can lead to multiplicative gain control in a single neuron by background inputs [5]. From a computational perspective, the sensitivity of the firing rate to mean or variance can be thought of as distinguishing the neuron's function as either an integrator (greater sensitivity to the mean) or a differentiator/coincidence detector (greater sensitivity to fluctuations, as quantified by the variance) [3,6,7].

An alternative method of characterizing a neuron's input-to-output transformation is through a linear/nonlinear (LN) cascade model [8,9]. These models comprise a set of linear filters or receptive field that selects particular features from the input; the filter output is transformed by a nonlinear threshold stage into a time-varying firing rate. Spike-triggered covariance analysis [10,11] reconstructs a model with multiple features from a neuron's input/output data. It has been widely employed to characterize both neural systems [12–15] and single neurons or neuron models subject to current or conductance inputs [16–19].

Generally, results of reverse correlation analysis may depend on the statistics of the stimulus used to sample the model [15,19–25]. While some of the dependence on stimulus statistics in the response of a neuron or neural system may reflect underlying plasticity, in some cases, the rapid timescale of the changes suggests the action of intrinsic nonlinearities in systems with *fixed* parameters [16,19,25–29], which changes the *effective* computation of a neuron.

Our goal here is to unify the *f-I* curve description of variance-dependent adaptive computation with that given by the LN model: we present analytical results showing that the variance-dependent modulation of the firing rate is closely related to adaptive changes in the *recovered* LN model if a fixed underlying model is assumed. When the model relies only on a single feature, we find that such a system can show only a single type of gain modulation, which accompanies an interesting asymptotic scaling behavior. With multiple features, the model can show more diverse adaptive behaviors, exemplified by two conductance-based models that we will study.

Results

Diverse Variance-Dependent Gain Modulations without Spike Rate Adaptation

Recently, Higgs et al. [3] and Arsiero et al. [4] identified different forms of variance-dependent change in the *f-I* curves of

Author Summary

Many neurons are known to achieve a wide dynamic range by adaptively changing their computational input/output function according to the input statistics. These adaptive changes can be very rapid, and it has been suggested that a component of this adaptation could be purely input-driven: even a fixed neural system can show apparent adaptive behavior since inputs with different statistics interact with the nonlinearity of the system in different ways. In this paper, we show how a single neuron's intrinsic computational function can dictate such input-driven changes in its response to varying input statistics, which begets a relationship between two different characterizations of neural function—in terms of mean firing rate and in terms of generating precise spike timing. We then apply our results to two biophysically defined model neurons, which have significantly different response patterns to inputs with various statistics. Our model of intrinsic adaptation explains their behaviors well. Contrary to the picture that neurons carry out a stereotyped computation on their inputs, our results show that even in the simplest cases they have simple yet effective mechanisms by which they can adapt to their input. Adaptation to stimulus statistics, therefore, is built into the most basic single neuron computations.

various neuron types in avian brainstem and in cortex. Depending on the type, neurons can have either increasing or decreasing gain in the *f-I* curve with increasing variance. These papers linked the phenomenon to mechanisms underlying spike rate adaptation, such as slow afterhyperpolarization (sAHP) currents and slow sodium channel inactivation. We recently showed [7] that a standard Hodgkin–Huxley (HH) neuron model, lacking spike rate adaptation, can show two different types of variance-dependent gain modulation simply by tuning the maximal conductance parameters of the model. These differences in gain modulation correspond to two different regimes in the space of conductance parameters. In one regime, which includes the standard parameters, a neuron periodically fires to a sufficiently large constant input current. In the other regime, a neuron never fires to a constant input regardless of its magnitude, but responds only to rapid fluctuations. This rarely discussed property has been termed *class 3 excitability* [30,31]. Higgs et al. [3] proposed that the type of gain modulation classifies the neuron as an integrator or differentiator.

Here, we examine two models that show these different forms of variance-dependent gain modulation without spike rate adaptation, and study the resulting LN models sampled with different stimulus statistics. We show that these *fixed* models generate variance-dependent gain modulation, and that this gain modulation is well predicted by aspects of the LN models derived from white noise stimulation. The two models are both based on the HH [32] active currents; one model is the standard HH model, and the other (HHLS) has lower Na^+ and higher K^+ conductances. The HHLS model is a class 3 neuron and responds only to a rapidly changing input. For this reason, the HHLS model can be thought of as behaving more like a differentiator than an integrator [3,7].

Figure 1 shows the different gain modulation behaviors of the HH and HHLS conductance-based models. For the HH model, Figure 1A, the *f-I* curves in the presence of noise are similar to the noiseless case except that they are increasingly smoothed at the threshold. In contrast, Figure 1C shows that the *f-I* curves of the HHLS model never converge toward each other as the noise level increases. This case resembles that of layer 5 pyramidal neurons in rat medial prefrontal cortex [4], as well as nucleus laminaris (NL) neurons in the chick auditory brainstem and some pyramidal neurons in layer 2/3 of rat neocortex [3]. While for these layer 2/3 neurons, there is evidence that this change in *f-I* curve slope may be related to the sAHP current [3], at steady state this effect can be obtained in general by tuning the maximal conductances without introducing any mechanism for spike rate adaptation [7].

Gain Modulation and Adaptation of Fixed Models

For a system described by an LN model with a single feature, we derive an equation relating the slopes of the firing rate with respect to stimulus mean and variance. We then consider gain modulation in a system with multiple relevant features and derive a series of equations relating gain change to properties of the spike-triggered average and spike-triggered covariance. Throughout, we assume that the underlying system is fixed, and that its parameter settings do not depend on stimulus statistics. For example, if the model has a single exponential filter with a time constant τ, we assume that τ does not change with the stimulus mean (I_0) or variance (σ^2). However, this does not mean that the model shows a single response pattern regardless of the statistical structure of stimuli. The sampled LN description of a nonlinear system with fixed parameters—even when the underlying model is an LN model [25]—can show interaction with the input statistics, leading to different LN model descriptions for different input parameters [19,25,27–29]. We refer to this as *intrinsic adaptation.*

One-Dimensional Model

An LN model is composed of its relevant features $\{\varepsilon_\mu(t)\}$ ($\mu = 1,2,\ldots,n$)), which act as linear filters on an incoming stimulus, and a probability to spike given the filtered stimulus, *P*(spike|filtered stimulus). For a Gaussian white noise stimulus with mean I_0 and variance σ^2, the firing rate is

$$f\left(I_0,\sigma^2\right)=\int \mathrm{d}\mathbf{x}\; P(\mathrm{spike}|I_0\bar{\boldsymbol{\varepsilon}}+\mathbf{x})p(\mathbf{x}) \qquad (1)$$

where $\bar{\boldsymbol{\varepsilon}}=\int_0^\infty \boldsymbol{\varepsilon}(\tau)\mathrm{d}\tau$ is the time-integrated filter and $\mathbf{x}$ is the mean-subtracted noise stimulus filtered by the *n* relevant features. $p(\mathbf{x})$ is an *n*-dimensional Gaussian distribution with variance σ^2. We refer to the Materials and Methods section for a more detailed account of the model.

For a one-dimensional model $n = 1$, Equation 1 can be rewritten with change of variables

$$f\left(I_0,\sigma^2\right)=\int_{-\infty}^{\infty} \mathrm{d}x\; P(\mathrm{spike}|x)p(x-I_0\bar{\varepsilon}) \qquad (2)$$

Since $p(x)$ is Gaussian, it is also the kernel or Green's function of a diffusion equation in terms of (x,σ^2) and therefore so is $p(x-I_0\bar{\varepsilon})$ in terms of (I_0,σ^2). In other words, we have

$$\left(\frac{\partial}{\partial\sigma^2}-\frac{1}{2}\frac{\partial^2}{\partial x^2}\right)p(x-I_0\bar{\varepsilon})$$
$$=\left(\frac{\partial}{\partial\sigma^2}-\frac{1}{2\bar{\varepsilon}^2}\frac{\partial^2}{\partial I_0^2}\right)p(x-I_0\bar{\varepsilon})=0$$

Now operating with $\left(\frac{\partial}{\partial\sigma^2}-\frac{1}{2\bar{\varepsilon}^2}\frac{\partial^2}{\partial I_0^2}\right)$ on both sides of the equation, $p(x-I_0\bar{\varepsilon})$ is the only term on the left hand side of Equation 2 that

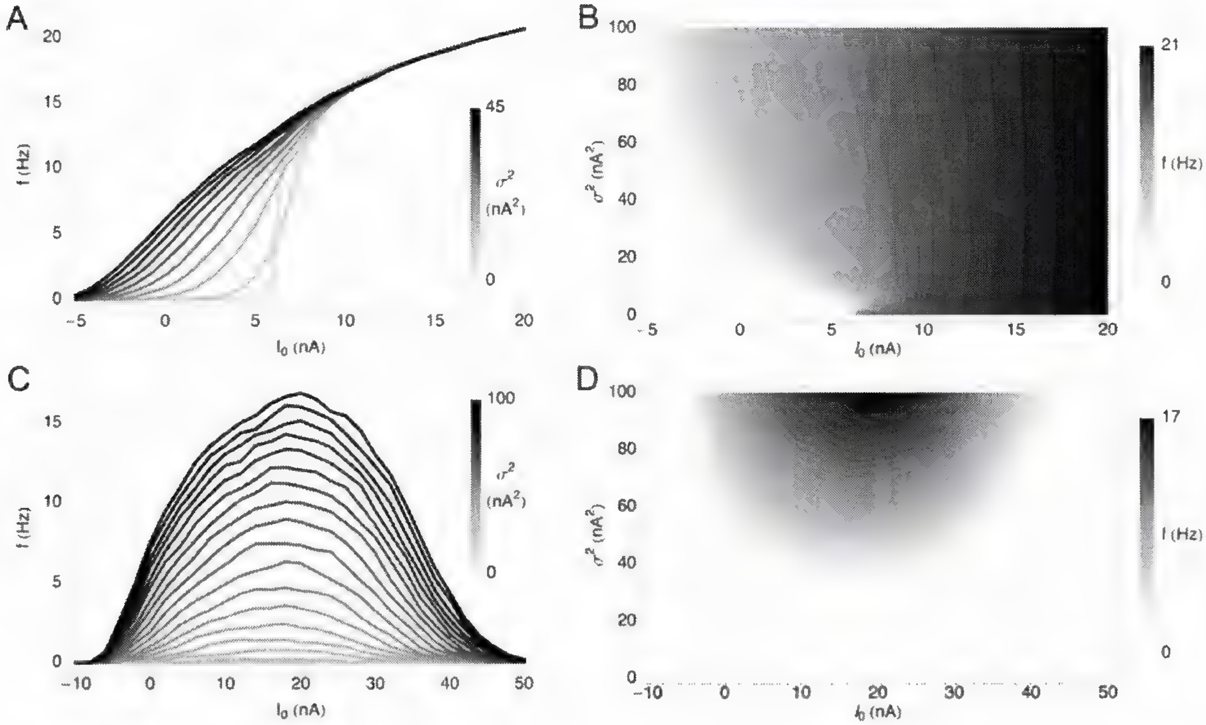


Figure 1. Variance-Dependent Gain Modulation of the HH and HHLS Model. Each model is simulated as described in the Materials and Methods section. (A) *f-I* curves of a standard HH model for differing 10 variances (σ^2) from 0 to 45 nA^2. The topmost trace is the response to the highest variance. Each curve is obtained with 31 mean values (I_0) ranging from −5 to 20 nA. (B) The same data as (A) plotted in the (mean, variance) plane. Lighter shades represent higher firing rates. We used cubic spline interpolation for points not included in the simulated data. (C,D) *f-I* curves of the HHLS model as in (A) and (B). 10 means from −10 to 50 nA and 21 variances from 0 to 100 nA^2 are used.
doi:10.1371/journal.pcbi.1000119.g001

depends on (I_0,σ^2) and therefore the right hand side of Equation 2 vanishes. Thus one finds

$$2\bar{\varepsilon}^2 \frac{\partial f}{\partial \sigma^2} = \frac{\partial^2 f}{\partial I_0^2} \qquad (3)$$

The boundary condition is given by evaluating Equation 2 as $\sigma^2 \rightarrow 0$; here the Gaussian distribution becomes a delta function

$$\lim_{\sigma^2 \rightarrow 0} p(x - I_0\bar{\varepsilon}) = \delta(x - I_0\bar{\varepsilon})$$

and the boundary condition is given by the zero-noise *f-I* curve. Thus, when a model depends only on a single feature, $\varepsilon(t)$, the *f-I* curve with a noisy input is given by a simple diffusion-like equation, Equation 3, with a single parameter, the time integrated filter, $\bar{\varepsilon} = \int_0^\infty \varepsilon(\tau)d\tau$, determining the diffusion constant $1/2\bar{\varepsilon}^2$.

Equation 3 states that the variance-dependent change in the firing rate is simply determined by the curvature of the *f-I* curve. Thus, a one-dimensional system displays only a single type of noise-induced gain modulation: as in diffusion, an *f-I* curve is gradually smoothed and flattened as the variance increases. Given a boundary condition, such as an *f-I* curve for a particular variance, the family of *f-I* relations can be reconstructed up to a scale factor by solving Equation 3. For example, one can predict how the neuron would respond to a noise stimulus based on its output in the absence of noise. Note that the solution of Equation 3 generalizes a classical result [33] based on a binary nonlinearity to a simple closed form which applies to any type of nonlinearity.

Figure 2A and 2B show a solution of Equation 3. While this one-dimensional model is based on the simplest and most general assumptions, it provides insights into the structure of variance-dependent gain modulation. The boundary condition is an *f-I* curve with no noise, $f = (I+0.1)^{1/2}$ for $I>0$ and $f = 0$ for $I\leq 0$, which imitates the general behavior of many dynamical neuron models around rheobase [34–36]. Compared with the HH conductance-based model, Equation 3 captures qualitative characteristics of the HH *f-I* curve despite differences due to the increased complexity of the HH model over a 1D LN model: in Figure 2A and 2B, there is a positive curvature (second derivative of firing rate with respect to current) of the *f-I* curve below rheobase related to the increase of the firing rate with increasing variance. In contrast, the behavior of the HHLS model cannot be described by Equation 3. Even though the *f-I* curves in Figure 1C mostly have negative curvature, the firing rate keeps increasing with variance, implying that the HHLS model cannot be described by a one-dimensional LN model.

We also compared Equation 3 with the *f-I* curves from two commonly used simple neuron models, the leaky integrate-and-fire (LIF) model (Figure 2C), and a similar model with minimal nonlinearity, the quadratic integrate-and-fire (QIF) model [37,38] (Figure 2D). The *f-I* curves of the two models are similar but have subtle differences: in the LIF model, firing rate never decreases with noise, even though parameters were chosen to induce a large negative curvature, as shown analytically in Text S1. The QIF

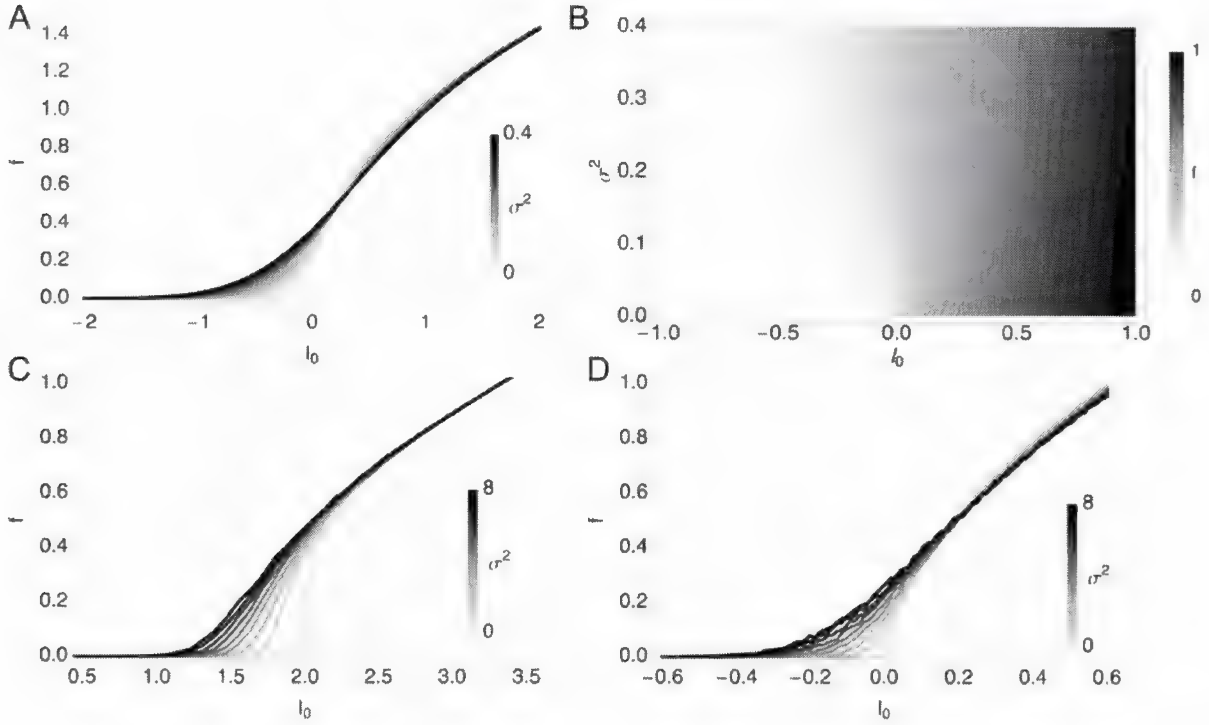


Figure 2. Variance-Dependent Gain Modulation of One-Dimensional Models. (A) Variance-dependent *f-I* curves of a one-dimensional model from the solution of Equation 3 with the boundary condition, $f = (I+0.1)^{1/2}$ for $I>0$ and $f=0$ for $I \leq 0$ at zero noise. (B) The firing rates of A in the (mean, variance) plane. (C) *f-I* curves of an LIF model. (D) *f-I* curves of a QIF model. The model parameters for the LIF and QIF are in the Materials and Methods section. We used 50 mean (I_0) values from 0 to 4 (LIF) and from −2 to 2 (QIF), and 8 variances (σ^2) from 0 to 8 for both models.

doi:10.1371/journal.pcbi.1000119.g002

model behavior is much more similar to the 1D LN model, marked by a slight decrease in firing rate at large I_0. From this perspective, the QIF is a *simpler* model in terms of the LN description despite the dynamical nonlinearity.

It is interesting to note that for one-dimensional models, the gain modulation given by Equation 3 depends only on the boundary condition, which implicitly describes how an input with a given mean samples the nonlinearity, but not explicitly on the details of filters or nonlinearity. An ideal differentiator, where firing rate is independent of the stimulus mean, is realized only when the filter has zero integral, $\bar{\varepsilon} = 0$. This is also the criterion that would be satisfied if the filter itself were ideally differentiating. We will return to the relationship between the LN model functional description and that of the *f-I* curves in the Discussion.

Multidimensional Models

Here we examine gain modulation in the case of a system with multiple relevant features. In this case, one cannot derive a single simple equation such as Equation 3. Instead, we derive relationships between the characteristics of $f(I_0,\sigma)$ curves and quantities calculated using white noise analysis.

Fixed multidimensional models can display far more complex response patterns to different stimulus statistics than one-dimensional models, because linear components in the model can now interact nonlinearly [29]. For example, in white noise analysis, as the stimulus variance increases, the distribution of the filtered stimuli also expands and probes different regions of the nonlinear threshold structure of the model. This induces a variance-dependent rotation among the filters recovered through sampling by white noise analysis, and the corresponding changes in the spike-triggered average, spike-triggered covariance, and the sampled nonlinearity [19].

Here, we relate parameters of the changing spike-triggered average and spike-triggered covariance description to the form of the *f-I* curves. The relationships are derived by taking derivatives of each side of Equation 1 with respect to I_0 and σ^2 (see Materials and Methods section). The first order in I_0 establishes the relationship between the STA and the gain of the *f-I* curve with respect to the mean

$$\frac{\partial \log f}{\partial I_0} = \frac{1}{\sigma^2}\overline{\mathrm{STA}}, \quad \overline{\mathrm{STA}} = \int_0^\infty \mathrm{d}\tau \mathrm{STA}(\tau) \qquad (4)$$

The second order leads to a relationship between the second derivative of the *f-I* curve and the covariance matrix

$$\frac{\partial^2 \log f}{\partial I_0^2} = \frac{1}{\sigma^4}\overline{\Delta C}, \quad \overline{\Delta C} = \int \mathrm{d}\tau \mathrm{d}\tau' \Delta C(\tau,\tau') \qquad (5)$$

The gain with respect to the variance is

$$\frac{\partial \log f}{\partial \sigma^2} = \frac{1}{2\sigma^4}\left(\mathrm{Tr}\Delta C + \|\mathrm{STA}\|^2\right) \qquad (6)$$

where

$$\mathrm{Tr}\Delta C = \int \mathrm{d}\tau \Delta C(\tau,\tau), \quad \|\mathrm{STA}\|^2 = \int \mathrm{d}\tau \mathrm{STA}(\tau)^2$$

Equations 4–6 show how the nonlinear gain of an *f-I* curve with respect to input mean and variance is related to intrinsic adaptation as observed through changes in the STA and STC. Note that Equations 4–6 apply to one-dimensional LN models as well. In that case, the STA has the same shape as the feature in the model, and only its magnitude varies according to the overlap integral, Equation 1, between the nonlinearity of the model and the prior stimulus. This is the same for the STC, and thus Equations 4–6 are not independent. This leads to a single form of variance gain modulation, given by Equation 3. However, in a multidimensional model, changing the stimulus mean shifts the nonlinearity in a single direction, $\overline{\text{STA}}$, while increasing the variance expands the prior in every direction in the stimulus space. Therefore, the overlap integral can show more diverse behaviors.

Conductance-Based Models

We now examine whether the gain modulation behaviors we have described can be captured by a multi-dimensional LN model. We tested this by computing *f-I* curves, spike-triggered averages and the spike-triggered covariance matrices for the noise-driven HH and HHLS models for a range of input statistics. Figure 3A, B, and C show the result of fitting simulation data from the HH (left) and HHLS (right) model to Equations 4, 5, and 6, respectively. The linear relationships are quite clear in Figure 3A

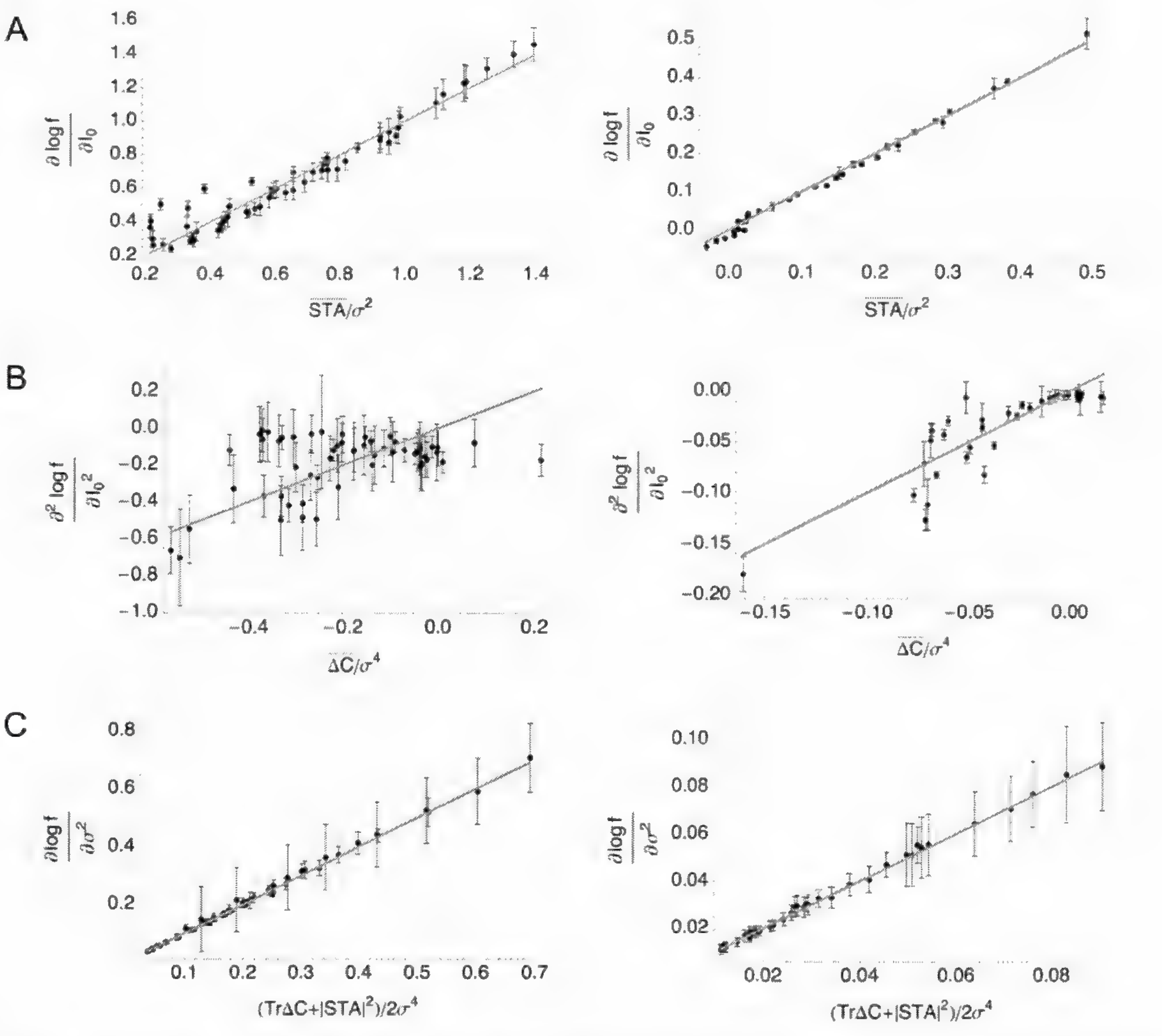


Figure 3. Derivatives of the Firing Rate Curves with Respect to Mean and Variance Related to Quantities Obtained by White Noise Analysis for the Standard HH (Left) and HHLS (Right) Models. Each point is calculated from the simulated data with a selected (mean, variance) input parameter pair, as described in the Materials and Methods section, and the gray lines represent our theoretical predictions, Equations 4–6, which hold when the variance dependent change in *f-I* curves is only due to intrinsic adaptation. (A) Gain versus the norm of the STA, as in Equation 4. (B) Gain change versus the spike-triggered covariance term of Equation 5. (C) Change of firing rate with respect to variance versus the function of the STA and spike-triggered covariance given in Equation 6.
doi:10.1371/journal.pcbi.1000119.g003

and 3C which show the gains with respect to mean and variance. Figure 3B involves the curvature of *f-I* curves, which is more difficult to calculate accurately, and shows larger errors. In every case, goodness of fit is $p<1.3\times10^{-6}$ and $p<5.8\times10^{-6}$ for the HH and HHLS where the upper bounds of *p*-values are given by the case of Equation 5, corresponding to Figure 3B. These results show that intrinsic adaptation of the LN model predicts the form of noise-induced gain modulation for these models.

Gain Rescaling of One-Dimensional Models

Here we discuss a consequence of intrinsic adaptation for neuronal encoding of mean and variance information for a one-dimensional model. In this case, Equation 3 completely specifies intrinsic adaptation, and therefore we will focus on this case.

Our first observation is that Equation 3 is invariant under the simultaneous rescaling of the mean and standard deviation, $I_0\rightarrow\alpha I_0$, $\sigma\rightarrow\alpha\sigma$, where α is an arbitrary positive number. This invariance is preserved if the solution is also a function of only a dimensionless variable I_0/σ, which would represent a signal-to-noise ratio if we describe the neuron's input/output function in terms of an *f-I* curve at a fixed noise level σ. Note that this situation is analogous to the Weber–Fechner [39,40] and Fitts' law [41], which states that perception tends to depend on only dimensionless variables that are invariant under scaling of the absolute magnitude of stimulus [42]. However, the invariance of Equation 3 under the scaling of a stimulus does not necessarily lead to the invariance of a firing rate solution. By rewriting Equation 2 in terms of the "rescaled" variables, $y=x/\sigma$ and $\mu=I_0/\sigma$, we get

$$f\left(\mu,\sigma^2\right)=\frac{1}{\sqrt{2\pi}}\int \mathrm{d}y\, \mathrm{e}^{-(y-\mu\bar{\varepsilon})^2/2} f_0\left(\frac{y\sigma}{\bar{\varepsilon}}\right) \qquad (7)$$

where $f_0(I)=P(\text{spike}\,|\,I\bar{\varepsilon})$ is an *f-I* curve with no noise. Thus, the scaling of $f(I_0,\sigma^2)$ with standard deviation depends on the boundary condition, $f_0(I)$, which in principle can be any arbitrary function.

Nevertheless, in practice, the *f-I* curves of many dynamical neurons are not completely arbitrary but can share a simple scaling property, at least asymptotically. For example, in the QIF and many other neuron models, the *f-I* curve with no noise asymptotically follows a power law $f_0\sim(I_0-I_c)^{1/2}$ around the rheobase I_c [34–36]. In general, if $f_0(I)\propto I^\alpha$ asymptotically in such a regime, from Equation 7, the firing rate is asymptotically factorized into a σ dependent and $\mu=I_0/\sigma$ dependent part as

$$f\left(\mu,\sigma^2\right)\propto\sigma^\alpha F(\mu),\quad F(\mu)=\frac{1}{\sqrt{2\pi}}\int \mathrm{d}y\, \mathrm{e}^{-(y-\mu\bar{\varepsilon})^2/2} y^\alpha \qquad (8)$$

In other words, I_0/σ becomes an *intermediate asymptotic* of the *f-I* curves [43].

To test to what extent this scaling relationship holds in the models we have considered, we calculated the *rescaled relative gain* of the *f-I* curves, which we define as $(\sigma/f)\,\partial f/\partial I_0=\sigma\,\partial\log f/\partial I_0$; the rescaled relative gain of Equation 8 depends only on $\mu=I_0/\sigma$, not on σ. Thus, if the rescaling strictly holds, this becomes a single-valued function of the signal-to-noise ratio, I_0/σ, regardless of the noise level σ.

We find evidence for this form of variance rescaling in the QIF, LIF, and HH models. Figure 4 shows the rescaled gains evaluated from the simulated data. The QIF and HH case, Figure 4B and 4D, match well with the solution of Equation 3, Figure 4A. In the LIF case, Figure 4C, the relative gain shows deviations at low variance, but it approaches a variance-independent limit at large σ. We also present an analytic account in Text S1. On the other hand, in Figure 4E, the HHLS model does not exhibit this form of asymptotic scaling at all. The role of the signal-to-noise ratio, I_0/σ, in the HHLS model appears to be quite distinct from the other models. In summary, Equation 3 predicts that one-dimensional LN models will have the tendency to decrease gain with increasing noise level. However, if the *f-I* curve of a neuron is power-law-like, the resulting gain modulation will be such that the neuron's sensitivity to mean stimulus change at various noise levels is governed only by the signal-to-noise ratio.

Discussion

In this paper, we have obtained analytical relationships between noise-dependent gain modulation of *f-I* curves and properties of the sampled linear/nonlinear model. We have shown that gain control arises as a simple consequence of the nonlinearity of the LN model, even with no changes in any underlying parameters.

For a system described by an LN model with only one relevant feature, a simple single-parameter diffusion relationship relates the *f-I* curves at different variances, where the role of the diffusion coefficient is taken by the integral of the STA. This form strictly limits the possible forms of gain modulation that may be manifested by such a system. The result qualitatively describes the variance dependent gain modulation of different neuron models such as the LIF, QIF, and standard HH neuron models. Models based on dynamical spike generation, such as QIF, showed better agreement with this result than the LIF model. The QIF model case is a good example of how a nonlinear dynamical system can be mapped onto an LN model description [19,44]. The QIF model has a single dynamical equation whose subthreshold dynamics are captured approximately by a linear kernel, which takes the role of the feature; one can then determine a threshold which acts as a binary decision boundary for spiking. Thus, it is reasonable that the QIF model and the one-dimensional LN model show a similar response pattern to a noisy input. When the system has multiple relevant features, we obtain equations relating the gain with respect to the input mean and the input variance to parameters of the STA and STC. We verified these results using HH neurons displaying two different forms of noise-induced gain control.

Previous work has related different gain control behaviors to a neuron's function as an integrator or a differentiator [3,7]. From an LN model perspective, the neuron's function is defined by specific properties of the filter or filters $\varepsilon(t)$. An integrating filter would consist of entirely positive weights; for leaky integrators these weights will decay at large negative times. A differentiating filter implements a local subtraction of the stimulus, and so should consist of a bimodal form where the positive weights approximately cancel the negative weights.

In general, characterizations of neural function by LN model and by *f-I* curves are quite distinct. The *f-I* approach we have discussed here describes the encoding of stationary statistical properties of the stimulus by time-averaged firing rate, while the LN model describes the encoding of specific input fluctuations by single spikes, generally under a particular choice of stimulus statistics. Indeed, the LN characterization can change with the driving stimulus distribution, even, in principle, from an integrator to a differentiator. Thus, a model may, for example, act as a differentiator on short timescales but as an integrator on longer timescales. For systems whose LN approximation varies with mean and variance, the neuron's effective computation changes with stimulus statistics, and so does the information that is represented.

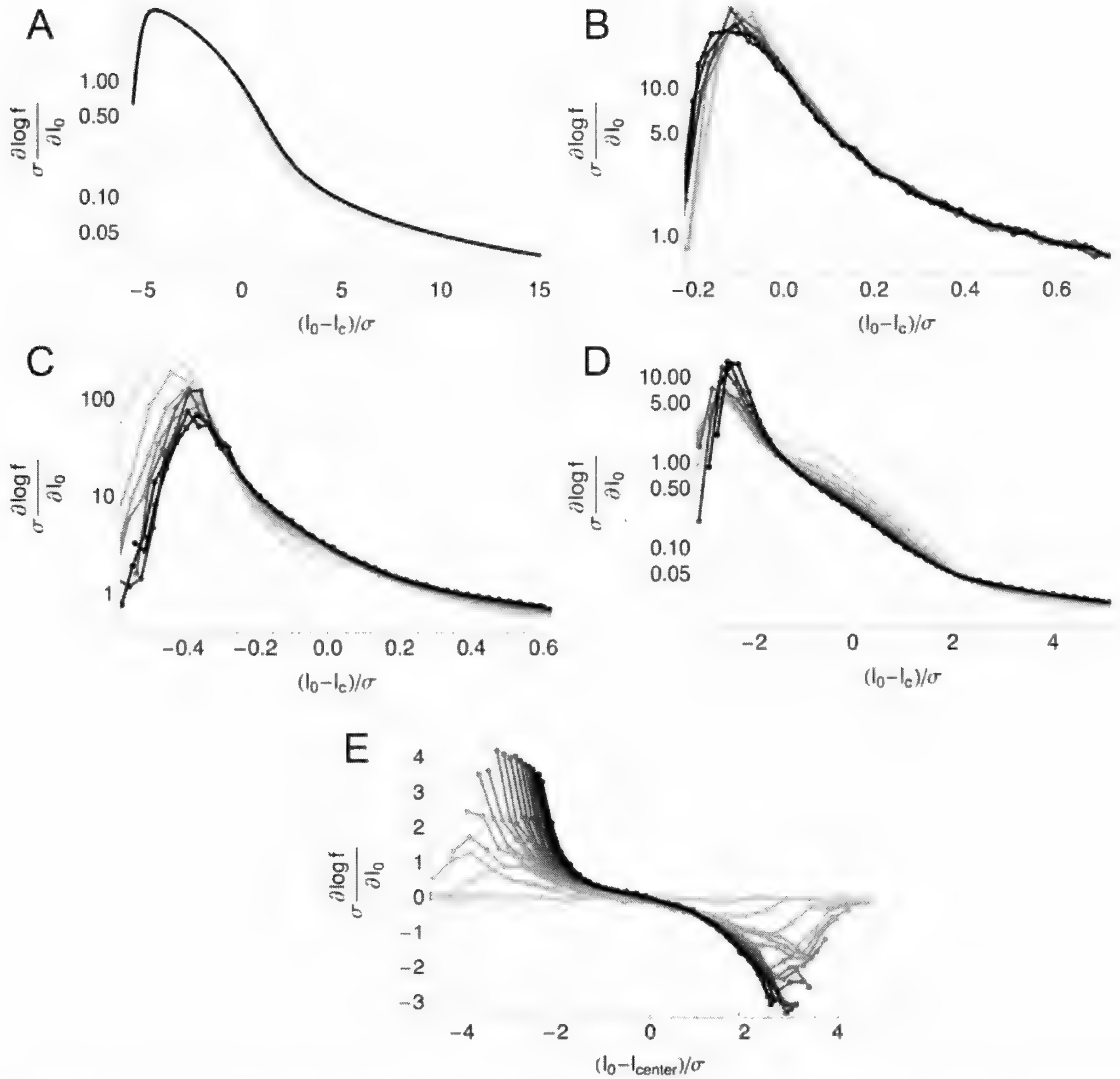


Figure 4. Rescaled Relative Gains of Variance-Dependent *f-I* Curves. (A) The one-dimensional LN, (B) QIF, and (C) LIF models. The same data as Figure 2 are used. (D) The standard HH model from Figure 1A and 1B. (E) The HHLS model from Figure 1C and 1D. Since the HHLS does not have a rheobase, we instead used $I_{center} = 20$ nA at which the variance-dependent firing rate increase is maximal.
doi:10.1371/journal.pcbi.1000119.g004

One might then ask how the system can decode the represented information. It has been proposed that statistics of the spike train might provide the information required to decode slower-varying stimulus parameters [22,45]. The possibility of distinguishing between responses to different stimulus statistics using the firing rate alone depends on the properties of the *f-I* curves.

The primary focus of this work is the restricted problem of single neurons responding to driving currents, where the integrated synaptic current in an in vivo-like condition is approximated to be a (filtered) Gaussian white noise [46–50]. However, our derivations can apply to arbitrary neural systems driven by white noise inputs, if *f-I* curves are interpreted as tuning functions with respect to the mean stimulus parameter. Given the generality of our results for neural systems, it would be interesting to test our results in cases where firing is driven by an external stimulus. A good candidate would be retinal ganglion cells, which are well-described by LN-type models [9,14,51–53], show adaptation to stimulus statistics on multiple timescales [23,54] and display a variety of dimensionalities in their feature space [14].

A limitation of the tests we have performed here is a restriction to the low firing rate regime where spike-triggered reverse correlation captures most of the dependence of firing probability on the stimulus. The effects of interspike interaction can be significant [16,17,55] and models with spike history feedback have been developed to account for this [44,51,56,57]. We have not investigated how spike history effects would impact our results.

Although evidence suggests that gain modulation by noise may be enhanced by slow afterhyperpolarization currents underlying spike frequency adaptation [3], these slow currents are not required to generate gain enhancement in simple neuron models [7,19,25–29]. While one may generate diverse types of noise-induced gain modulation only by modifying the mechanism of generating a spike independent of spike history [7], in realistic situations, slow adaptation currents are present and will affect neural responses over many timescales [58–60]. In principle, it is possible to extend our result to include these effects: *f-I* curves under conditions of spike frequency adaptation have been already discussed [61–63] and can be compared to LN models with spike history feedback. However, our goal here was to demonstrate the effects that can occur independent of slow adaptation currents and before such currents have acted to shift neuronal coding properties.

The suggestive form of our result for one-dimensional LN models led us to look for a representation of neuronal output that is invariant under change in the input noise level. Our motivation is based on a simple principle of dimensional analysis: the gains of the *f-I* curves with noise may be asymptotically described by a signal-to-noise ratio, a dimensionless variable depending only on the stimulus itself. We showed that this may occur if the *f-I* curve with no noise obeys asymptotic power-law properties. Such a property has been determined to arise both from the bifurcation patterns of spike generation [31,34,35] and due to spike rate adaptation [61]. This relationship implies that the gain of the firing rate as a function of the mean should scale inversely with the standard deviation. Scaling of the gain of the nonlinear decision function with the stimulus standard deviation has been observed to some degree in a number of neural systems [10,15,22–25,29,64–67]. Such scaling guarantees maximal transmission of information [10,22]. As we and others have proposed, a static model might suffice to explain this phenomenon [25,27], although in some cases slow adaptation currents are known to contribute [65,66].

In summary, we have presented theoretically derived relationships between the variance-dependent gain modulation of *f-I* curves and intrinsic adaptation in neural coding. In real neural systems, any type of gain modulation likely results from many different mechanisms, possibly involving long-time scale dynamics. Our results show that observed forms of gain modulation may be a result of a pre-existing static nonlinearity that reacts to changes in the stimulus statistics robustly and almost instantaneously.

Materials and Methods

Biophysical Models

We used two single compartmental models with Hodgkin-Huxley (HH) active currents. The first one is an HH model with standard parameters while the second model (HHLS) has a lower Na^+ and higher K^+ maximal conductance. The voltage changes are described by [32]

$$C\frac{\mathrm{d}V}{\mathrm{d}t} = -\bar{g}_L(V-E_L) - \bar{g}_{Na}m^3h(V-E_{Na}) - \bar{g}_K n^4(V-E_K) + I(t)$$

and the activation variables *m*, *n*, and *h* behave according to

$$\tau_z(V)\frac{\mathrm{d}z}{\mathrm{d}t} = \bar{z}(V) - z, \quad \tau_z = \frac{1}{\alpha_z+\beta_z}, \quad \bar{z} = \frac{\alpha_z}{\alpha_z+\beta_z}, \quad z = m,n,h$$

where

$$\alpha_m = \frac{0.1(V+40)}{1-\exp[-0.1(V+40)]}, \quad \beta_m = 4\exp[-0.0556(V+65)],$$

$$\alpha_h = 0.07\exp[0.05(V+65)], \quad \beta_h = \frac{1}{1+\exp[-0.1(V+35)]},$$

$$\alpha_n = \frac{0.01(V+55)}{1-\exp[-0.1(V+55)]}, \quad \beta_n = 0.125\exp[-0.0125(V+65)]$$

The voltage V is in millivolts (mV).

For the HH model, the conductance parameters are $\bar{g}_K = 36$ mS/cm^2 and $\bar{g}_{Na} = 120$ mS/cm^2. The HHLS model has $\bar{g}_K = 41$ mS/cm^2 and $\bar{g}_{Na} = 79$ mS/cm^2. All other parameters are common to both models. The leak conductance is $\bar{g}_L = 0.3$ mS/cm^2 and the membrane capacitance per area C is 1 μF/cm^2. The reversal potentials are $E_L = -54.3$ mV, $E_{Na} = 50$ mV, and $E_K = -77$ mV. The membrane area is 10^{-3} cm^2, so that a current density of 1 μA/cm^2 corresponds to a current of 1 nA.

All simulations of these models were done with the NEURON simulation environment [68]. Gaussian white noise currents with various means and variances are generated with an update rate of 5 kHz ($\mathrm{d}t = 0.2$ ms) and delivered into the model via current clamp. For the *f-I* curves, we simulated 4 min of input for each mean and variance pair. The whole procedure was repeated five times to estimate the variance of the *f-I* relationship, σ_{repeat}.

We ran another set of simulations for reverse correlation analysis and collected about 100,000 spikes for each stimulus condition. The means and variances of the Gaussian noisy stimuli were chosen such that the mean firing rate did not exceed 10 Hz, and we selected eight means and seven variances for the HH model, and nine means and four variances for the HHLS model.

Integrate-and-Fire-Type Models

In addition to the conductance-based model, we investigated the behavior of two heuristic model neurons driven by a noisy current input. Each model consists of a single dynamical equation describing voltage fluctuations of the form

$$C\frac{\mathrm{d}V}{\mathrm{d}t} = L(V) + I(t)$$

The first model is a leaky integrate-and-fire (LIF) model [69,70], for which $L(V) = -g_L(V-E_L)$. We used the parameters $g_L = 2$, $E_L = 0$, and $C = 1$. Since this choice of $L(V)$ cannot generate a spike, we additionally imposed a spiking threshold $V_{th} = 1$ and reset voltage $V_{\mathrm{reset}} = -3$.

The second is a quadratic integrate-and-fire (QIF) model [31,37,38], for which $L(V) = g_L(V-E_L)(V-V_{th})/\Delta V$ where $\Delta V = V_{th} - E_L > 0$. We used $g_L = 0.5$, $E_L = 0$, $V_{th} = 0.1$, and $C = 1$. In this model, the voltage V can increase without bound; such a trajectory is defined to be a spike if it crosses $V_{\mathrm{spike}} = 5$. After spiking, the system is reset to $V_{\mathrm{reset}} = 0$.

These two models are simulated using a fourth-order Runge-Kutta integration method with an integration time step of $\mathrm{d}t = 0.01$. The input current $I(t)$ was Gaussian white noise, updated at each time step, with a range of means and variances. The *f-I* curves were obtained from 1,000 s of stimulation for each (mean,variance) condition. We then compared the *f-I* curves from these models with the relationship derived in the Results section, Equation 5. A numerical solution of the partial differential equation was obtained using a PDE solver in Mathematica (Wolfram Research, Inc.).

Linear/Nonlinear Model

We use the linear/nonlinear (LN) cascade model framework to describe a neuron's input/output relation. We will focus on the dependence of the firing rate of a fixed LN model on the mean and variance of a Gaussian white noise input.

We will take the driving input to be $I(t) = I_0 + \xi(t)$ where I_0 is the mean and $\xi(t)$ is a Gaussian white noise with variance σ^2 and zero mean. The linear part of the model selects, by linear filtering, a subset of the possible stimuli probed by $I(t)$. That subset is expressed as n relevant features $\{\varepsilon_\mu(t)\}$, $(\mu = 1,2,\ldots,n)$. Interpreted as vectors, the components of any stimulus that are relevant to changing the firing rate can be expressed in terms of projections onto these features. The firing rate of the model for a given temporal sequence $I(t)$ depends only on $\mathbf{s}$, the input filtered by the n relevant features. Thus the firing rate from the given stimulus depends on the convolution of the input with all n features and can be written as $P(\text{spike} \mid \mathbf{s} = I_0\bar{\varepsilon}+\mathbf{x})$ where

$$\bar{\varepsilon}_\mu = \int_0^\infty \mathrm{d}\tau\, \varepsilon_\mu(\tau), \quad x_\mu = \int_0^\infty \mathrm{d}\tau\, \varepsilon_\mu(\tau)\xi(t-\tau)$$

Since $I(t)$ is white noise with stationary statistics, the projections x_μ can be taken to be stationary random variables chosen from a Gaussian distribution at each t.

Given the filtered stimulus, a nonlinear decision function $P(\text{spike} \mid I_0\bar{\varepsilon}+\mathbf{x})$ generates the instantaneous time-varying firing rate. For a specified model and stimulus statistics, the mean firing rate $f(I_0,\sigma^2) = P(\text{spike})$ is simply

$$f\left(I_0,\sigma^2\right) = \int \mathrm{d}\mathbf{s}\; P(\text{spike}|\mathbf{s})P(\mathbf{s}) = \int \mathrm{d}\mathbf{x}\; P(\text{spike}|I_0\bar{\varepsilon}+\mathbf{x})p(\mathbf{x}) \qquad (9)$$

where

$$p(\mathbf{x}) = \frac{1}{\left(2\pi\sigma^2\right)^{n/2}} \cdot \exp\left(-\frac{1}{2\sigma^2}\|\mathbf{x}\|^2\right)$$

Equation 9 describes an *f-I* curve of the model in the presence of added noise with variance σ^2. The slope or *gain* of the firing rate with respect to mean or variance can be computed if $P(\text{spike} \mid I_0\bar{\varepsilon}+\mathbf{x})$ is known. However, the gains can be also obtained in terms of the first and second moments of $P(\text{spike} \mid I_0\bar{\varepsilon}+\mathbf{x})$, which can be measured *directly* by reverse correlation analysis.

Reverse Correlation Analysis

We used spike-triggered reverse correlation to probe the computation of the model neurons through an LN model. We collected about 100,000 spikes and corresponding ensembles of spike triggered stimulus histories in a 30 ms long time window preceding each spike.

From the spike-triggered input ensembles, we calculated spike-triggered averages (STAs) and spike-triggered covariances (STCs). The STA is simply the average of the set of stimuli that led to spikes subtracted from the mean of the "prior" stimulus distribution, the distribution of all stimuli independent of spiking output

$$\mathrm{STA}(t) = \langle I\left(t_{\text{spike}}-t\right)\rangle_{\text{spike}} - \langle I\rangle_{\text{prior}} = \langle \xi\left(t_{\text{spike}}-t\right)\rangle_{\text{spike}} \qquad (10)$$

Therefore, one may consider only the noise part of the zero mean stimulus.

When computing the STC, the prior's covariance is subtracted

$$\begin{aligned}\Delta C(t,t') &= C_{\text{spike}} - C_{\text{prior}} \\ &= \langle\{\xi(t_{\text{spike}}-t) - \mathrm{STA}(t)\} \\ &\quad \{\xi(t_{\text{spike}}-t') - \mathrm{STA}(t')\}\rangle_{\text{spike}} - C_{\text{prior}}\end{aligned} \qquad (11)$$

Statistical Analysis

In calculating the slope and curvature of the *f-I* curves, we used 6–10 degree polynomial fitting of the *f-I* curves, where in any single case, the lowest degree was used which provided both a good fit and smoothness. From the fitting procedure, we obtained the standard deviation of the residuals, σ_{fit}. This was repeated five times for *f-I* curves computed using different noise samples, and from this we computed σ_{repeat}, the standard deviation of each computed slope and curvature. We estimated the total error of our calculation as $\sigma_{\text{total}} = (\sigma_{\text{repeat}}{}^2 + \sigma_{\text{fit}}{}^2)^{1/2}$. In practice, σ_{repeat} was always greater than σ_{fit} by an order of magnitude. This σ_{total} was used for the error bars in Figure 3.

To evaluate the goodness of fit in Figure 3, we used the Pearson χ^2 test by using the reduced χ^2 statistic

$$\chi^2 = \sum \frac{(O-E)^2}{\sigma_{\text{total}}^2}$$

where O and E represent the right and left hand sides of Equations 4–6, respectively. From this, the p-values are estimated from the cumulative density function of the χ^2 distribution, $Q(\chi^2/k,k)$. The degree of freedom is $k = 54$ and $k = 34$ for the HH and HHLS, respectively.

Derivation of Equations 4–6

We first present two key identities: the first one, which depends on the form of $\mathbf{s}$ having additive mean and noise components, is a change of variables for the gradient of $P(\text{spike} \mid \mathbf{x}+I_0\bar{\varepsilon})$

$$\frac{\partial P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})}{\partial I_0} = \sum_\mu \bar{\varepsilon}_\mu \frac{\partial P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})}{\partial x_\mu} \qquad (12)$$

Secondly, when x is a Gaussian random variable with zero mean and variance σ^2, by using integration by parts in can be seen that any function $F(x)$ satisfies

$$\langle F'(x)\rangle = \frac{1}{\sigma^2}\langle xF(x)\rangle \qquad (13)$$

$$\begin{aligned}\langle F''(x)\rangle &= \frac{1}{\sigma^2}\langle [xF(x)]'\rangle - \frac{1}{\sigma^2}\langle F(x)\rangle \\ &= \frac{1}{\sigma^4}\langle x^2F(x)\rangle - \frac{1}{\sigma^2}\langle F(x)\rangle\end{aligned}$$

Then, we first take derivatives of both sides of Equation 9 (or equivalently Equation 1), by I_0 and σ^2, and apply Equations 12 and 13. The first order in I_0 is

$$\begin{aligned}\frac{\partial \log f}{\partial I_0} &= \frac{1}{f}\frac{\partial f}{\partial I_0} = \frac{1}{f}\sum_{\mu}\bar{\varepsilon}_{\mu}\langle\frac{\partial}{\partial x_{\mu}}P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})\rangle_{\mathbf{x}} \\ &= \frac{1}{\sigma^2}\cdot\frac{1}{f}\sum_{\mu}\bar{\varepsilon}_{\mu}\langle x_{\mu}P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})\rangle_{\mathbf{x}}\end{aligned} \tag{14}$$

The second order is given by

$$\begin{aligned}\frac{\partial^2 \log f}{\partial I_0^2} &= \frac{1}{f}\frac{\partial^2 f}{\partial I_0^2} - \frac{1}{f^2}\left(\frac{\partial f}{\partial I_0}\right)^2, \\ \frac{\partial^2 f}{\partial I_0^2} &= \sum_{\mu,\nu}\bar{\varepsilon}_{\mu}\bar{\varepsilon}_{\nu}\langle\frac{\partial}{\partial x_{\mu}}\frac{\partial}{\partial x_{\nu}}P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})\rangle_{\mathbf{x}} \\ &= \frac{1}{\sigma^4}\sum_{\mu,\nu}\bar{\varepsilon}_{\mu}\bar{\varepsilon}_{\nu}\langle\left(x_{\mu}x_{\nu}-\sigma^2\delta_{\mu\nu}\right)P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})\rangle_{\mathbf{x}},\end{aligned} \tag{15}$$

where $\delta_{\mu\nu}$ is a Kronecker delta symbol. The gain with respect to variance is

$$\begin{aligned}\frac{\partial f}{\partial \sigma^2} &= -\frac{n}{2\sigma^2}f + \frac{1}{2\sigma^4}\sum_{\mu}\langle x_{\mu}^2 P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})\rangle_{x} \\ &= \frac{1}{2\sigma^4}\sum_{\mu}\langle\left(x_{\mu}^2-\sigma^2\right)P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})\rangle_{\mathbf{x}}\end{aligned} \tag{16}$$

Now, we show how the right hand sides of Equations 14–16 correspond to the STA and the STC. Given a Gaussian white noise signal $\xi(t)$, we can split it as $\xi=\xi_{\parallel}+\xi_{\perp}$, where $\xi_{\parallel}$ belongs to the space spanned by our basis features $\{\varepsilon_{\mu}\}$, and therefore relevant to spiking. $\xi_{\perp}$ is the orthogonal or irrelevant part. $\xi_{\parallel}$ can be written as

$$\xi_{\parallel}(t) = \boldsymbol{\varepsilon}\cdot\mathbf{x} = \sum_{\mu}x_{\mu}\varepsilon_{\mu}, \quad x_{\mu} = \int_0^{\infty} \mathrm{d}\tau\,\varepsilon_{\mu}(\tau)\xi(t-\tau)$$

Again, $\mathbf{x}$ is a Gaussian variable from a distribution Equation 9.

The STA is

$$\text{STA} = \langle\xi\rangle_{\text{spike}} = \langle\xi_{\parallel}\rangle_{\text{spike}} = \int \mathrm{d}^n\mathbf{x}(\varepsilon\cdot\mathbf{x})P(\mathbf{x}+I_0\bar{\varepsilon}|\text{spike})$$

since $\xi_{\perp}$ is irrelevant and does not make any contribution. Here we use Bayes theorem

$$\frac{P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})}{P(\text{spike})} = \frac{P(\mathbf{x}+I_0\bar{\varepsilon}|\text{spike})}{P(\mathbf{x}+I_0\bar{\varepsilon})}$$

As in Equation 9, $P(\mathbf{s}=\mathbf{x}+I_0\bar{\varepsilon})=p(\mathbf{x})$, and therefore the STA becomes

$$\begin{aligned}\text{STA} &= \int \mathrm{d}^n\mathbf{x}(\varepsilon\cdot\mathbf{x})\frac{P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})}{P(\text{spike})}p(\mathbf{x}) \\ &= \frac{1}{f}\sum_{\mu}\varepsilon_{\mu}\langle x_{\mu}P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})\rangle_{\mathbf{x}}\end{aligned}$$

Comparing this result with Equation 14, we obtain Equation 4.

A similar calculation for the second order [19] shows

$$\begin{aligned}\Delta C(t,t') = &\frac{1}{f}\sum_{\mu,\nu}\varepsilon_{\mu}(t)\varepsilon_{\nu}(t')\langle\left(x_{\mu}x_{\nu}-\sigma^2\delta_{\mu\nu}\right)P(\text{spike}|\mathbf{x}+I_0\bar{\varepsilon})\rangle_{\mathbf{x}} \\ &-\text{STA}(t)\cdot\text{STA}(t')\end{aligned}$$

This result, combined with Equations 15 and 16, leads to Equations 5 and 6, respectively.

Supporting Information

Text S1. Firing Rate of the LIF Model with Noisy Stimuli. Found at: doi:10.1371/journal.pcbi.1000119.s001 (0.09 MB DOC)

Author Contributions

Conceived and designed the experiments: SH. Analyzed the data: SH. Contributed reagents/materials/analysis tools: BL. Wrote the paper: SH BL AF. Derived the equations: SH.

References

1. Chance FS, Abbott LF, Reyes AD (2002) Gain modulation from background synaptic input. Neuron 35: 773–782.
2. Fellous JM, Rudolph M, Destexhe A, Sejnowski TJ (2003) Synaptic background noise controls the input/output characteristics of single cells in an in vitro model of in vivo activity. Neuroscience 122: 811–829.
3. Higgs MH, Slee SJ, Spain WJ (2006) Diversity of gain modulation by noise in neocortical neurons: regulation by the slow afterhyperpolarization conductance. J Neurosci 26: 8787–8799.
4. Arsiero M, Lüscher HR, Lundstrom BN, Giugliano M (2007) The impact of input fluctuations on the frequency–current relationships of layer 5 pyramidal neurons in the rat medial prefrontal cortex. J Neurosci 27: 3274–3284.
5. Prescott SA, Koninck YD (2003) Gain control of firing rate by shunting inhibition: roles of synaptic noise and dendritic saturation. Proc Natl Acad Sci USA 100: 2076–2081.
6. Prescott SA, Ratté S, De Koninck Y, Sejnowski TJ (2006) Nonlinear interaction between shunting and adaptation controls a switch between integration and coincidence detection in pyramidal neurons. J Neurosci 26: 9084–9097.
7. Lundstrom BN, Hong S, Fairhall AL (2008) Two computational regimes of a single-compartment neuron separated by a planar boundary in conductance space. Neural Comput 20: 1239–1260.
8. Victor J, Shapley R (1980) A method of nonlinear analysis in the frequency domain. Biophys J 29: 459–483.
9. Meister M, Berry II MJ (1999) The neural code of the retina. Neuron 22: 435–450.
10. Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive rescaling maximizes information transmission. Neuron 26: 695–702.
11. Simoncelli EP, Paninski L, Pillow J, Schwartz O (2004) Characterization of neural responses with stochastic stimuli. In: Gazzaniga M, ed. The Cognitive Neurosciences. 3rd edition. Cambridge (Massachusetts): MIT Press.
12. Rust NC, Schwartz O, Movshon JA, Simoncelli EP (2005) Spatiotemporal elements of macaque V1 receptive fields. Neuron 46: 945–956.
13. Stanley GB, Lei FF, Dan Y (1999) Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. J Neurosci 19: 8036–8042.
14. Fairhall AL, Burlingame CA, Narasimhan R, Harris RA, Puchalla JL, et al. (2006) Selectivity for multiple stimulus features in retinal ganglion cells. J Neurophysiol 96: 2724–2738.
15. Maravall M, Petersen RS, Fairhall AL, Arabzadeh E, Diamond ME (2007) Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. PLoS Biol 5: e19. doi:10.1371/journal.pbio.0050019.
16. Agüera y Arcas B, Fairhall AL (2003) What causes a neuron to spike? Neural Comput 15: 1715–1749.
17. Agüera y Arcas B, Fairhall AL, Bialek W (2003) Computation in a single neuron: Hodgkin and Huxley revisited. Neural Comput 15: 1789–1807.

18. Slee SJ, Higgs MH, Fairhall AL, Spain WJ (2005) Two-dimensional time coding in the auditory brainstem. J Neurosci 25: 9978–9988.
19. Hong S, Agüera y Arcas B, Fairhall AL (2007) Single neuron computation: from dynamical system to feature detector. Neural Comput 19: 3133–3172.
20. Atick JJ (1992) Could information theory provide an ecological theory of sensory processing? Network (Bristol, England) 3: 213–251.
21. Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J Neurosci 20: 2315–2331.
22. Fairhall A, Lewen G, Bialek W, de Ruyter van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. Nature 412: 787–792.
23. Baccus SA, Meister M (2002) Fast and slow contrast adaptation in retinal circuitry. Neuron 36: 909–919.
24. Nagel KI, Doupe AJ (2006) Temporal processing and adaptation in the songbird auditory forebrain. Neuron 51: 845–859.
25. Gaudry KS, Reinagel P (2007) Benefits of contrast normalization demonstrated in neurons and model cells. J Neurosci 27: 8071–8079.
26. Rudd ME, Brown LG (1997) Noise adaptation in integrate-and-fire neurons. Neural Comput 9: 1047–1069.
27. Paninski L, Lau B, Reyes AD (2003) Noise-driven adaptation: in vitro and mathematical analysis. Neurocomputing 52: 877–883.
28. Yu Y, Lee TS (2003) Dynamical mechanisms underlying contrast gain control in single neurons. Phys Rev E 68: 011901.
29. Borst A, Flanagin VL, Sompolinsky H (2005) Adaptation without parameter change: dynamic gain control in motion detection. Proc Natl Acad Sci USA 102: 6172–6176.
30. Hodgkin AL (1948) The local electric changes associated with repetitive action in a non-medullated axon. J Physiol 107: 165–181.
31. Izhikevich EM (2006) Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting. Cambridge (Massachusetts): MIT Press.
32. Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol 463: 391–407.
33. Spekreijse H, Reits D (1982) Sequential analysis of the visual evoked potential system in man: nonlinear analysis of a sandwich system. Ann N Y Acad Sci 388: 72–97.
34. Ermentrout GB (1994) Reduction of conductance-based models with slow synapses to neural nets. Neural Comput 6: 679–695.
35. Rinzel JM, Ermentrout GB (1989) Analysis of neuronal excitability. In: Koch C, Segev I, eds. Methods in Neuronal Modeling: From Synapses to Networks. Cambridge (Massachusetts): MIT Press. pp 135–170.
36. Hoppensteadt F, Izhikevich EM (1997) Weakly Connected Neural Nets. Berlin: Springer-Verlag.
37. Ermentrout GB, Kopell N (1986) Parabolic bursting in an excitable system coupled with a slow oscillation. SIAM J Appl Math 4: 233–253.
38. Ermentrout B (1996) Type I membranes, phase resetting curves, and synchrony. Neural Comput 8: 979–1001.
39. Weber EH (1834) De Pulsu, Resorptione, Auditu et Tactu. Annotiones Anatomicae et Physiologicae. Lipsiae: Koehler.
40. Fechner G (1966) Elements of Psychophysics. New York: Holt, Rinehart and Winston.
41. Fitts PM (1954) The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol 47: 381–391.
42. Stevens SS (1986) Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects. Piscataway (New Jersey): Transaction Publishers.
43. Barenblatt GI (2003) Scaling. Cambridge, UK: Cambridge University Press.
44. Gerstner W, Kistler W (2002) Spiking Neuron Models: Single Neurons, Populations, Plasticity. Cambridge, UK: Cambridge University Press.
45. Lundstrom BN, Fairhall AL (2006) Decoding stimulus variance from a distributional neural code of interspike intervals. J Neurosci 26: 9030–9037.
46. Gerstein GL, Mandelbrot B (1964) Random walk models for the spike activity of a single neuron. Biophys J 4: 41–68.
47. Bryant HL, Segundo JP (1976) Spike initiation by transmembrane current: a white-noise analysis. J Physiol 260: 279–314.
48. Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. Science 268: 1503–1506.
49. Destexhe A, Paré D (1999) Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo. J Neurophysiol 81: 1531–1547.
50. Rudolph M, Destexhe A (2003) Characterization of subthreshold voltage fluctuations in neuronal membranes. Neural Comput 15: 2577–2618.
51. Keat J, Reinagel P, Reid RC, Meister M (2001) Predicting every spike: a model for the responses of visual neurons. Neuron 30: 803–817.
52. Chichilnisky EJ (2001) A simple white noise analysis of neuronal light responses. Network (Bristol, England) 12: 199–213.
53. Pillow JW, Paninski L, Uzzell VJ, Simoncelli EP, Chichilnisky EJ (2005) Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. J Neurosci 25: 11003–11013.
54. Smirnakis SM, Berry MJ, Warland DK, Bialek W, Meister M (1997) Adaptation of retinal processing to image contrast and spatial scale. Nature 386: 69–73.
55. Pillow JW, Simoncelli EP (2003) Biases in white noise analysis due to non-Poisson spike generation. Neurocomputing 52–54: 109–115.
56. Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN (2005) A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. J Neurophysiol 93: 1074–1089.
57. Paninski L, Pillow J, Lewi J (2006) Statistical models for neural encoding, decoding, and optimal stimulus design. Prog Brain Res 165: 493–507.
58. Schwindt PC, Spain WJ, Foehring RC, Stafstrom CE, Chubb MC, et al. (1988) Multiple potassium conductances and their functions in neurons from cat sensorimotor cortex in vitro. J Neurophysiol 59: 424–449.
59. Spain WJ, Schwindt PC, Crill WE (1991) Two transient potassium currents in layer V pyramidal neurones from cat sensorimotor cortex. J Physiol 434: 591–607.
60. La Camera G, Rauch A, Thurbon D, Lüscher HR, Senn W, et al. (2006) Multiple time scales of temporal response in pyramidal and fast spiking cortical neurons. J Neurophysiol 96: 3448–3464.
61. Ermentrout B (1998) Linearization of F-I curves by adaptation. Neural Comput 10: 1721–1729.
62. Benda J, Herz AVM (2003) A universal model for spike-frequency adaptation. Neural Comput 15: 2523–2564.
63. La Camera G, Rauch A, Lüscher HR, Senn W, Fusi S (2004) Minimal models of adapted neuronal response to in vivo-like input currents. Neural Comput 16: 2101–2124.
64. Kim KJ, Rieke F (2001) Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. J Neurosci 21: 287–299.
65. Arganda S, Guantes R, de Polavieja GG (2007) Sodium pumps adapt spike bursting to stimulus statistics. Nat Neurosci 10: 1467–1473.
66. Diaz-Quesada M, Maravall M (2008) Intrinsic mechanisms for adaptive gain rescaling in barrel cortex. J Neurosci 28: 696–710.
67. Ringach DL, Malone BJ (2007) The operating point of the cortex: neurons as large deviation detectors. J Neurosci 27: 7673–7683.
68. Hines ML, Carnevale NT (1997) The NEURON simulation environment. Neural Comput 9: 1179–1209.
69. Knight BW (1972) Dynamics of encoding in a population of neurons. J Gen Physiol 59: 734.
70. Tuckwell HC (1988) Introduction to Theoretical Neurobiology. Cambridge, UK: Cambridge University Press.

LETTER Communicated by Peter Rowat

Two Computational Regimes of a Single-Compartment Neuron Separated by a Planar Boundary in Conductance Space

Brian Nils Lundstrom
lundbr@u.washington.edu
Sungho Hong
shhong@u.washington.edu
Department of Physiology and Biophysics, University of Washington, Seattle, WA 98195, U.S.A.

Matthew H. Higgs
higgsm@u.washington.edu
Department of Physiology and Biophysics, University of Washington, Seattle, WA 98195, and Veterans Affairs Puget Sound Health System, Seattle, WA 98108, U.S.A.

Adrienne L. Fairhall
fairhall@u.washington.edu
Department of Physiology and Biophysics, University of Washington, Seattle, WA 98195, U.S.A.

Recent in vitro data show that neurons respond to input variance with varying sensitivities. Here we demonstrate that Hodgkin-Huxley (HH) neurons can operate in two computational regimes: one that is more sensitive to input variance (differentiating) and one that is less sensitive (integrating). A boundary plane in the 3D conductance space separates these two regimes. For a reduced HH model, this plane can be derived analytically from the V nullcline, thus suggesting a means of relating biophysical parameters to neural computation by analyzing the neuron's dynamical system.

1 Introduction

A neuron's ion channel configuration determines how it processes information. However, the relationship between a particular set of ion channels and the specific computation it performs is still not clear, even for basic Hodgkin-Huxley (HH) model neurons (Hodgkin & Huxley, 1952; Aguera y Arcas, Fairhall, & Bialek, 2003; Yu & Lee, 2003) and simplified versions of the model (Kepler, Abbott, & Marder, 1992; Gerstner & Kistler, 2002; Murray, 2002). In this letter, we examine information processing in the HH neuron, where this model neuron is taken to be a generic case with typical

Neural Computation **20, 1239–1260** (2008) © 2008 Massachusetts Institute of Technology

spike-triggering ion channels. We focus on the response of the space-clamped HH model to a time-varying synaptic current input.

We approximate the net synaptic current input at the soma as exponentially filtered gaussian white noise current (Gerstein & Mandelbrot, 1964; Bryant & Segundo, 1976; Tuckwell, 1988; Mainen & Sejnowski, 1995; Destexhe, Rudolph, Fellous, & Sejnowski, 2001; Rauch, La Camera, Luscher, Senn, & Fusi, 2003; Rudolph & Destexhe, 2003a; Richardson & Gerstner, 2005), in which the input mean reflects the average number of inputs, while fluctuations, quantified by the variance, generally depend on the degree of neuronal input synchrony (Destexhe & Paré, 1999; Moreno, de la Rocha, Renart, & Parga, 2002; Richardson, 2004). We use the firing rate-current (*f-I*) function as a straightforward means of examining the model's input-output response properties. If firing rate varies most strongly with mean current, we term the neuron an *integrator*; if the firing rate is sensitive to variance and relatively insensitive to the mean current, the neuron is fluctuation driven, and we refer to it as a *differentiator* (Abeles, 1982; Konig, Engel, & Singer, 1996; Higgs, Slee, & Spain, 2006). A differentiating neuron is thus characterized by a low or zero firing rate in response to a constant or zero variance input.

Naturally, a spectrum exists between these two classifications. In addition, the operating mode of a given neuron often depends on the mean input current. Many neurons function as differentiators at low, subthreshold mean currents but behave more like integrators at suprathreshold mean currents; other neurons differentiate their input at saturating mean currents. However, the purest neural differentiators, such as specialized coincidence detectors in the auditory brain stem (Oertel, 1983; Reyes, Rubel, & Spain, 1994; Rothman & Manis, 2003), appear entirely incapable of integration and never fire repetitively at any level of constant current stimulation.

Modeling studies suggest that neurons can function as either integrators or coincidence detectors (differentiators) based on the nature of incoming input (Gutkin, Ermentrout, & Rudolph, 2003; Rudolph & Destexhe, 2003b). In vitro studies suggest that high conductance inputs facilitate coincidence detection or differentiation (Gonsalves & Paller, 2000; Destexhe, Rudolph, & Paré, 2003; Prescott, Ratte, De Koninck, & Sejnowski, 2006), and at least one in vivo study suggests that synchronized inputs more effectively drive cortical spiking (Roy & Alloway, 2001). Recent in vitro studies suggest that neuronal firing rates are affected by the statistical properties of approximated synaptic inputs (Chance, Abbott, & Reyes, 2002; Fellous, Rudolph, Destexhe, & Sejnowski, 2003) and that different populations of neurons respond to input variance with differing sensitivities (Higgs et al., 2006; Arsiero, Luscher, Lundstrom, & Giugliano, 2007).

Our aim was to determine the biophysical parameter regime for which the HH neuron retains sensitivity to input variance even at high input means. We conclude that the HH neuron can process information in two fundamental ways, as both an integrator and a differentiator, and that despite the highly nonlinear nature of the model, a planar boundary in the space of maximal conductances separates these two regimes. Further, using a 2D simplification of the HH model, we find that this planar boundary can be derived directly from model equations and related dynamical system properties, thus demonstrating a simple link between the observed categories of computational function and the biophysical conductance parameters.

2 Hodgkin-Huxley Neuron as Both Integrator and Differentiator

Recent studies show that *f-I* curves from in vitro cortical pyramidal cells, hippocampal CAI pyramidal cells, and neurons from avian auditory brain stem demonstrate changes in sensitivity to input variance as the mean is increased and suggest that these changes are related to slow adaptation currents (Higgs et al., 2006; Prescott et al., 2006) or slow sodium inactivation (Arsiero et al., 2007). However, many simple neuron models, including the standard HH model neuron, the Connor-Stevens model neuron, and the Traub-Miles model neuron, show decreased sensitivity to input variance as the mean increases. We have found that for several simple model neurons, lowering the maximal sodium conductance G_{Na}, such as through voltage-dependent slow sodium inactivation, leads to an increased sensitivity to input variance at high means, as would be the case for differentiating neurons (see Figure 1). In this differentiator operating regime, these neurons do not fire repetitively to constant, or zero-variance, inputs.

Here, we consider only static maximal conductance values rather than any activity-dependent (LeMasson, Marder, & Abbott, 1993; Giugliano, Bove, & Grattarola, 1999) or voltage-dependent change. The membrane potential (V) of the space-clamped HH neuron is described by

$$C\frac{dV}{dt} = -G_{\mathrm{Na}}m^3h(V-E_{\mathrm{Na}}) - G_{\mathrm{K}}n^4(V-E_{\mathrm{K}}) - G_{\mathrm{Leak}}(V - E_{\mathrm{Leak}}) + I, \tag{2.1}$$

where C is capacitance, G_i are the maximal channel conductances, E_i are the reversal potentials, I is the external, injected current, and $[m, h, n]$ are the channel gating variables, which obey first-order dynamics (see appendix A). If we take this standard HH neuron and systematically lower its maximal sodium conductance, its response to inputs, as evaluated by an *f-I* curve, changes (see Figure 2). When G_{Na} is high, the HH neuron behaves more like an integrator, while when G_{Na} is low, the neuron behaves more like a differentiator.

Conceptually, these differences in the *f-I* curves (see Figure 1b, gray and black) result from changes in an effective voltage threshold for spiking, where spike initiation occurs when enough sodium channels open to counter outward currents. The maximal conductance G_{Na}, for example, approximately establishes the spiking threshold. Additionally, the availability

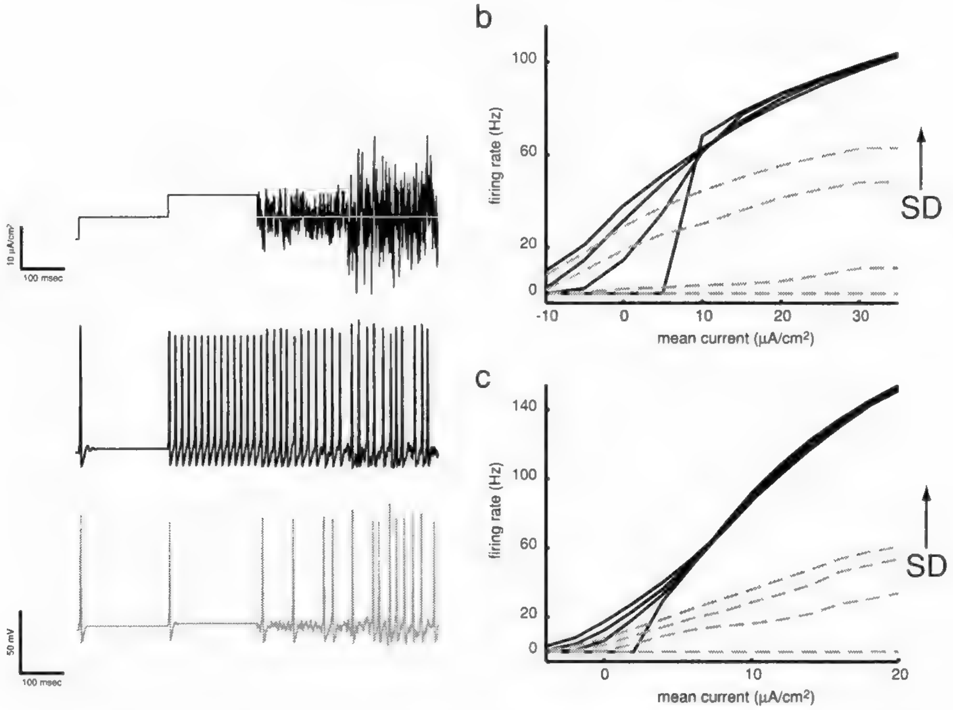


Figure 1: Spectrum of integration and differentiation. (a) Simple differentiators do not respond to zero variance inputs at steady state. This nonzero current stimulus undergoes a twofold increase in mean, [5 10] μA/cm^2, followed by a twofold increase in standard deviation, [2.5 5] μA/cm^2, and is presented to two different neuron models. Although not ideal, the black trace has integrator characteristics, whereby changes in standard deviation result in only very small firing rate changes. In contrast, firing rates in the gray trace increase as standard deviation increases, but the steady-state firing rate is always zero if input variance is zero. Decreased sodium conductance leads to neurons that are better differentiators in (b) the standard Hodgkin-Huxley model neuron and in (c) a neocortical model neuron (Gerstner & Kistler, 2002). Data are plotted for the Hodgkin-Huxley and neocortical models with high (black, solid lines) and low (gray, dotted lines) sodium conductances. Lines represent mean firing rate at steady state in response to input current; increased input standard deviation generally leads to increased firing rate. Models with low sodium conductance do not fire in response to zero-variance inputs at steady state. Parameters for *b*: HH: $G_{\mathrm{Na}} = [120\ 82]$ mS/cm^2, SD = [0 2 4 6] μA/cm^2. Parameters for *c*: cortical: $G_{\mathrm{Na}} = [50\ 25]$ mS/cm^2, SD = [0 2 3 4] μA/cm^2 neurons. Standard parameters (see appendix A) were used unless otherwise noted.

of sodium channels depends on the voltage-dependent inactivation variable h, adjusting the effective threshold value. Differentiation then occurs when slow components of the input are below this threshold, and firing results only from input noise. In certain regimes, such as when G_{Na} has been lowered sufficiently, this effective threshold is never surpassed as input mean

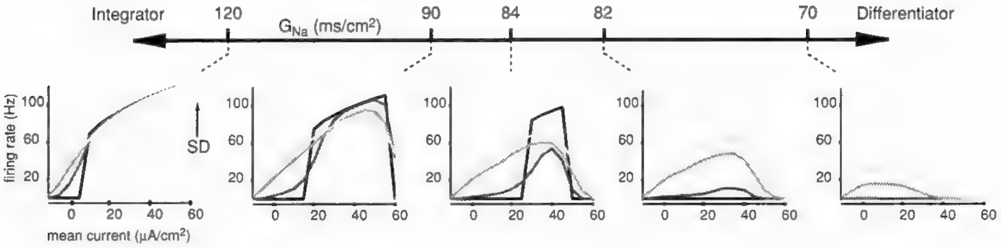


Figure 2: The HH neuron behaves more like a differentiator as its sodium conductance G_{Na} is lowered. As G_{Na} is lowered below ~83 mS/cm^2, the HH neuron ceases to respond to zero-variance input. Traces on each *f-I* plot from black to light gray represent increasing SD, where SD = [0 2 4 6] μA/cm^2. The first and fourth *f-I* plots from the left show the same data as in Figure 1b. Standard HH values (G_{Na}, G_{K}, G_{Leak}) = (120, 36, 0.3) mS/cm^2 were used unless otherwise noted (see appendix A).

increases; the voltage-dependent threshold keeps increasing slightly as the input mean increases until eventually the neuron is unable to spike. This effect, known as depolarization block, is observed for large inputs in the *f-I* curves of Figure 2. This conceptual explanation will be further developed in terms of the neuron's dynamical system.

3 A Planar Boundary Separates Integration from Differentiation

We simulated the HH model using a set of 56 different conductance values (G_{Na}, G_{K}, G_{Leak}) and from these determined *f-I* curves from each condition. As seen in Figure 2, at high G_{Na}, the HH neuron responds well to noiseless inputs, whereas at low G_{Na}, the neuron does not respond to zero-variance inputs. This transition from noiseless firing to no response is relatively sudden along the range of G_{Na} in which firing occurs. This transition can also be observed when G_{K} or G_{Leak} is increased or when these three conductances are modulated in linear combination according to[1]

$$G_{\mathrm{Na}} - 2.07G_{\mathrm{K}} - 22.8G_{\mathrm{Leak}} = 0, \tag{3.1}$$

---

[1] These coefficients are the result of a multiple least-squares regression of 56 conductance sets (G_{Na}, G_{K}, G_{Leak}) determined via simulation. For a given G_{Na} and G_{Leak}, G_{K} would initially take a relatively low value for which the neuron would fire to some mean input current I. G_{K} was then increased incrementally and *f-I* curves computed until the neuron would not fire to any mean input I; this was then one of the boundary sets. As the ratios of the conductances (i.e., $N = G_{\mathrm{Na}}/G_{\mathrm{Leak}}$ and $K = G_{\mathrm{K}}/G_{\mathrm{Leak}}$) increased, the steady-state voltages V_0 at the boundary current value (which we later define as I^*) generally decreased: mean –52.8; SD 2.0; range [−48.6 –54.8]. Conductances were chosen according to the following constraints: N was within the range [50 500], $G_{\mathrm{Na}} > 50$ mS/cm^2, G_{Leak} = 0.3, 1, or 2. The best fit of these data passes through or very close to the origin (see appendix B).

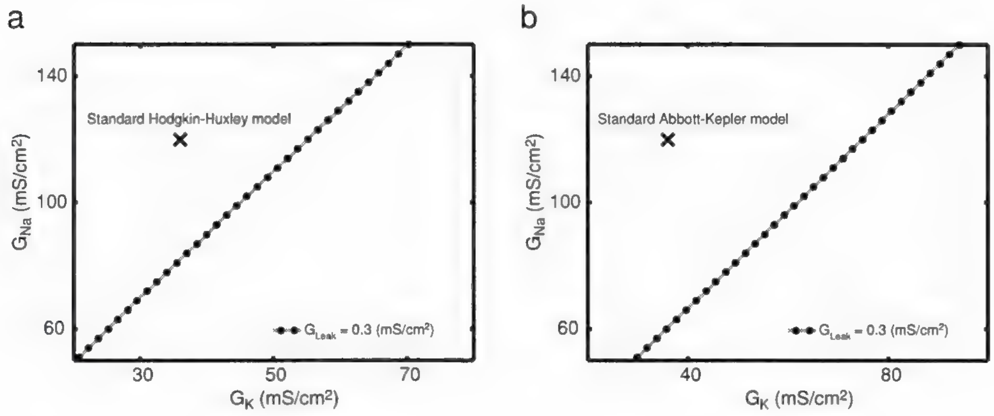


Figure 3: The boundary between no response to constant stimuli (differentiation) and firing to constant stimuli (integration) is described by a line in the two-dimensional space of G_{Na} and G_{K}. This boundary is similar for both the Hodgkin-Huxley (a) and Abbott-Kepler (b) neuron models. As G_{Leak} increases, the boundary remains linear with the same slope but an increasing intercept in G_{Na}. The black X's designate the standard parameters for the Hodgkin-Huxley and Abbott-Kepler models.

where conductances have units mS/cm^2. When this sum is less than zero, the HH neuron will not fire to a noiseless current, whereas if the sum is greater than zero, the neuron fires repetitively. Equation 3.1 describes the empirically observed plane that passes through the origin in the three-dimensional space of HH conductances, which is shown as a line for fixed G_{Leak} in Figure 3a. Here, if the neuron does not fire to a noiseless current, no matter how large the mean, then the neuron is considered to behave more like a differentiator.[2] In terms of the dynamical system, when the right-hand side of equation 3.2 is greater than zero, the system undergoes a Hopf bifurcation at a particular mean current (e.g., approximately 6.5 μA/cm^2 in Figure 1b, bottom black trace); however, if the right-hand side is less than zero, the one fixed point of the dynamical system is stable for all mean current values (as in Figures 5b and 5c).

Having obtained equation 3.1 through simulation, we would like to derive it directly from model equations. However, due to its four differential equations, the dynamical system of the HH model is difficult to analyze. Therefore, we use the two-dimensional Abbott-Kepler (AK) model, which is derived from the HH model equations and maintains a direct correspondence of parameters (Abbott & Kepler, 1990; Kepler et al., 1992). In the AK model, the four variables of the HH model (V, m, h, and n) are divided by timescale into two groups that become V and U, a fast activation and a slow recovery variable, respectively. The 2D AK model equations have the form (Abbott & Kepler, 1990; Kepler et al., 1992; Hong, Aguera y Arcas, & Fairhall, 2007):

---

[2] Although for the HH model this is an appropriate criterion, other models can show nonzero firing rates in response to zero-variance stimuli, while their *f-I* curves are differentiator-like, as in Figure 2. Further, even neurons that are here classified as integrators, as seen on the left side of Figure 2, display differentiator-like characteristics when mean input currents are low.

$$C\frac{dV}{dt} = -G_{\mathrm{Na}}m_{\infty}^{3}(V)h_{\infty}(U)(V - E_{\mathrm{Na}}) - G_{\mathrm{K}}n_{\infty}^{4}(U)(V - E_{\mathrm{K}}) \\ -G_{\mathrm{Leak}}(V - E_{\mathrm{Leak}}) + I, \tag{3.2}$$

$$\frac{dU}{dt} \\ = \frac{G_{\mathrm{Na}}(V - E_{\mathrm{Na}})m^{3}(V)(h(V) - h(U))/\tau_{h}(V) + 4G_{\mathrm{K}}(V - E_{\mathrm{K}})n^{3}(U)(n(V) - n(U))/\tau_{n}(V)}{G_{\mathrm{Na}}(V - E_{\mathrm{Na}})m^{3}(V)h'(U) + 4G_{\mathrm{K}}(V - E_{\mathrm{K}})n^{3}(U)n'(U)}. \tag{3.3}$$

Equation 3.2 has a similar form to equation 2.1, where the dynamics of the gating variables have been replaced by their steady-state voltage-dependent values m_{∞}, n_{∞}, and h_{∞}. Like the HH model, this model shows a boundary plane between integration and differentiation, which is described by[3]

$$G_{\mathrm{Na}} - 1.54G_{\mathrm{K}} - 16.7G_{\mathrm{Leak}} = 0, \tag{3.4}$$

where conductances have units $\mathrm{mS/cm^2}$. The line that is described by equation 3.4 when $G_{\mathrm{Leak}} = 0.3\ \mathrm{mS/cm^2}$ is shown in Figure 3b.

4 Deriving the Planar Boundary

The advantage of working with a two-dimensional dynamical system is that we can more easily understand how the empirical boundary equation, equation 3.4, emerges from the dynamical system, equations 3.2 and 3.3. We begin from the phase portrait of the AK model. The flow is determined by the V and U nullclines, defined as the set of points for which $dV/dt = 0$ and $dU/dt = 0$, respectively. The unique fixed point of the system is the point of intersection of the two nullclines (Gerstner & Kistler, 2002;

---

[3] This process was the same as for the HH model for 55 conductance sets. Again, as conductance ratios ($N = G_{\mathrm{Na}}/G_{\mathrm{Leak}}$ and $K = G_{\mathrm{K}}/G_{\mathrm{Leak}}$) increased, steady-state voltages V_0 at the boundary current value (which we will later define as I^*) decreased: (cf. Figure 6) mean −50.3; SD 0.8; range [−48.2 –51.2]. The best fit of these data passes through or very close to the origin (see appendix B).

Murray, 2002; Izhikevich, 2007). In the AK model, from equation 3.3, the U nullcline is given by $U = V$. The V nullcline generally takes an N-shape for intrinsically spiking neural models (Izhikevich, 2007) and is given by the solution of

$$\begin{aligned} 0 = &-G_{\mathrm{Na}} m_\infty^3(V) h_\infty(U)(V - E_{\mathrm{Na}}) - G_{\mathrm{K}} n_\infty^4(U)(V - E_{\mathrm{K}}) \\ &-G_{\mathrm{Leak}}(V - E_{\mathrm{Leak}}) + I, \end{aligned} \tag{4.1}$$

which for simplicity we express as

$$f(U, V) = I, \tag{4.2}$$

where

$$\begin{aligned} f(U, V) \equiv\, &G_{\mathrm{Na}} m_\infty^3(V) h_\infty(U)(V - E_{\mathrm{Na}}) + G_{\mathrm{K}} n_\infty^4(U)(V - E_{\mathrm{K}}) \\ &+ G_{\mathrm{Leak}}(V - E_{\mathrm{Leak}}). \end{aligned}$$

Thus, for the AK model, the shape of the V nullcline in (V, U) space depends nonlinearly on mean current I as well as on the parameters of equation 4.2. We will focus on the three maximal conductances (G_{Na}, G_{K}, G_{Leak}).

For any given set of the four parameters (G_{Na}, G_{K}, G_{Leak}, I), whether the neuron fires in response to DC, or noiseless, current is determined by the shape of the V nullcline and the position of the fixed point on this nullcline. If the V nullcline has an N-like shape, as seen in Figures 4a and 4b, the system is excitable. However, for some values of I, the V nullcline can flatten, and the system becomes unable to create a spike; this corresponds to depolarization block and is seen in the top contour of Figure 5a, which is the V nullcline for $I = 40\,\mu\mathrm{A/cm}^2$ and the given conductances. When the system is able to spike, the second factor determining DC response is where the fixed point, the intersection of the V nullcline with the line $U = V$, lies with respect to the local minimum of the V nullcline, or the left "knee" of the curve. When the fixed point is to the right of the knee, the neuron fires repetitively to zero-variance mean current (see Figures 4a and 4c), but when the fixed point is to the left of the local minimum, it does not (see Figures 4b and 4d).[4]

When the current is noisy, the position of the fixed point with respect to the minimum determines whether spike initiation is largely dictated by either the internal dynamics of the system (see Figure 4c) or by momentary

---

[4] This may not be strictly true. As has been noted for the Fitzhugh-Nagumo model (Izhikevich, 2007), loss of stability by the fixed point may not occur precisely at the local minimum in the V-nullcline, as can be verified by linearization and local stability analysis. The effect is small in the AK model (see appendix B).

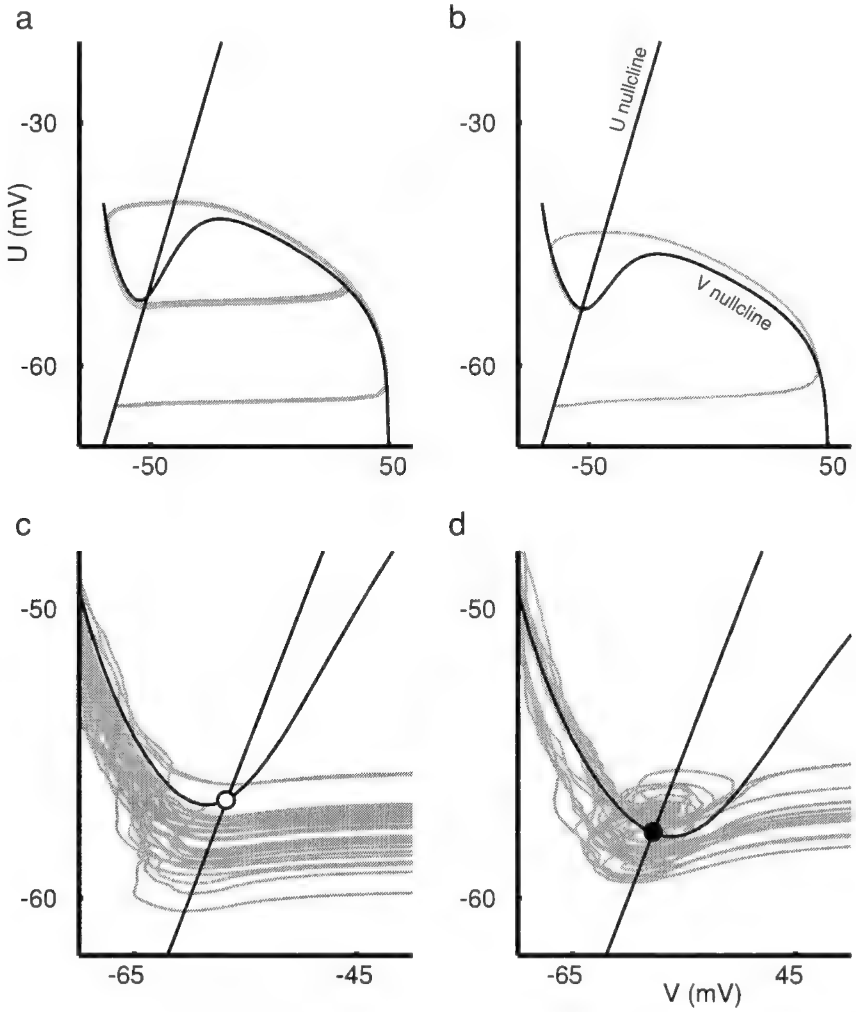


Figure 4: Phase portraits of the Abbott-Kepler neuron model. (a) The neuron fires repetitively to a mean current of 50 $\mu A/cm^2$ when sodium conductance is high ($G_{Na} = 120\ mS/cm^2$) such that the conductance values lie above the plane of equation 3.4 but (b) does not fire repetitively when the sodium conductance is low ($G_{Na} = 50\ mS/cm^2$). The straight black line is the U nullcline corresponding to $dU/dt = 0$, where U is a general recovery variable, while the curved black line is the V nullcline, satisfying $dV/dt = 0$, where V is the membrane voltage. The gray lines are the trajectories of (V, U) in time, where the loops in *a* are the limit cycles of spikes. For a mean of 20 $\mu A/cm^2$ and SD of 4 $\mu A/cm^2$, spike initiation is dominated by (c) the internal dynamics when G_{Na} is high but (d) when G_{Na} is low, spikes are initiated by input variance. The fixed points in the two cases are (c) unstable, open circle, and (d) stable, closed circle, respectively. Other model parameters include $G_K = 36\ mS/cm^2$ and $G_{Leak} = 0.3\ mS/cm^2$.

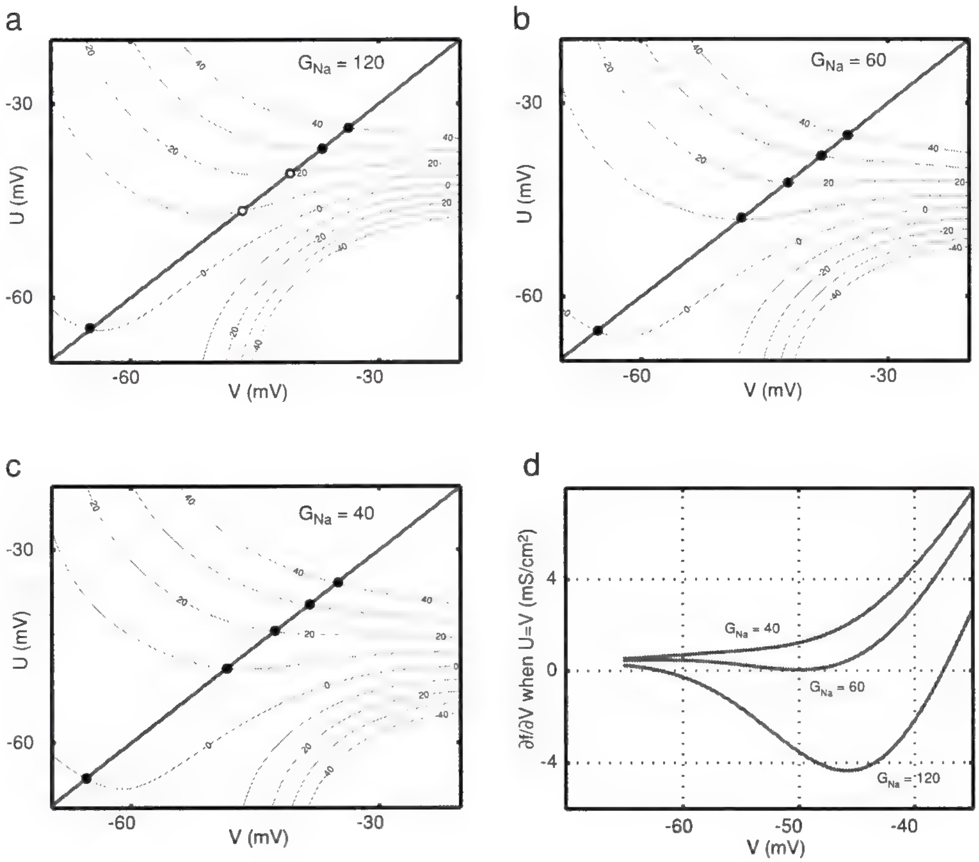


Figure 5: V nullclines described by $f(U, V) = I$. (a) The surface of equation 4.2 is plotted as a contour in current I (μA/cm^2) for the conductances (G_{Na}, G_{K}, G_{Leak}) = (120, 36, 0.3). (b) $f(U, V)$ for (G_{Na}, G_{K}, G_{Leak}) = (60, 36, 0.3), which is very close to the boundary. (c) $f(U, V)$ for (G_{Na}, G_{K}, G_{Leak}) = (40, 36, 0.3). Fixed points fall on the line $U = V$, where stable and unstable fixed points are represented by closed and open circles, respectively. (d) The partial derivative of $f(U, V)$ with respect to V evaluated for $U = V$ for three sets of conductances. The bottom ($G_{\text{Na}} = 120$ mS/cm^2) and top ($G_{\text{Na}} = 40$ mS/cm^2) traces represent approximately integrating and approximately differentiating neurons, respectively. The middle ($G_{\text{Na}} = 60$ mS/cm^2) trace lies approximately at the boundary between the two regimes. For this AK model, (G_{K}, G_{Leak}) were (36, 0.3) mS/cm^2.

large current fluctuations (see Figure 4d). These fluctuations cause leftward or rightward excursions along the V-axis, which may cause the neuron to cross an upward-curving threshold function situated slightly beneath, but approximately parallel to, the V nullcline (Hong et al., 2007). If a change in stimulus is gradual, the nullclines alter their geometry, which may prevent threshold crossing despite an equally large or larger stimulus perturbation.

Each point in the parameter space (G_{Na}, G_{K}, G_{Leak}, I) maps onto one of three possible conditions: no minimum exists in the V nullcline and thus the system is not excitable; or the system is excitable, with a fixed point to either the right or the left of the minimum. For the AK model to implement a perfect differentiator, it should not fire spontaneously for any mean current level I. For this to be the case, for values of I such that the V nullcline shows a minimum, the nullclines must have the configuration of Figure 4b or 4d. Thus, as I increases, the V-nullcline minimum will disappear before the fixed point shifts to the right of the V nullcline minimum. We use this to derive the boundary condition on conductance parameters.

Let us consider an example of the behavior of the system. In Figure 5a, we show contours of the surface, equation 4.2, which are a set of V nullclines each for different I. As I increases, the nullclines lose their N-like shape, and eventually the neuron is no longer excitable and undergoes depolarization block. For differentiating neurons, the line $U = V$ always intersects the V nullclines at a place where the V nullcline is decreasing; that is, the derivative of the V nullcline evaluated at the fixed point is negative. As can be seen in Figure 5, this means that $\partial f/\partial V > 0$. A boundary conductance (G_{Na}, G_{K}, G_{Leak}) gives, then, a set of I-dependent V nullclines for which this partial derivative has a minimum value of zero, as in Figure 5d when $G_{\mathrm{Na}} = 60\ \mathrm{mS/cm^2}$.

For the fixed point to occur to the left of the minimum, we require that $dU/dV < 0$ for all fixed points ($U_0(I)$, $V_0(I)$). Thus, the condition is given by

$$\left.\frac{dU}{dV}\right|_{U_0,V_0} = -\left.\frac{\partial f/\partial V}{\partial f/\partial U}\right|_{U_0,V_0} = 0. \tag{4.3}$$

One of the two fixed-point conditions is that all fixed points, independent of I, must satisfy $U = V$. Thus, fixed points for all I fall somewhere along the line $U = V$. Since we want to characterize the system's behavior for all I, we will thus evaluate the condition at $U = V$. Further, we will consider only the numerator of equation 4.3 (see also appendix B), giving the condition

$$\left.\frac{\partial f(U,V)}{\partial V}\right|_{U=V} = 0. \tag{4.4}$$

Figure 5d shows some examples of the function $\partial f/\partial V$, evaluated for several values of the conductances. Our boundary criteria require not only that the derivative is zero, but that zero is the minimal value for any I. Thus, at the point ($V_0(I^*)$, $U_0(I^*)$), where the derivative takes the value zero, the function $\partial f/\partial V$ must be at a minimum. Hence, we require that for some value of V, $V^* = V_0(I^*)$, where V^* is a function of I and the conductance parameters, equation 4.4 and the following hold simultaneously:

$$\frac{d}{dV}\left(\left.\frac{\partial f(U,V)}{\partial V}\right|_{U=V}\right) = 0. \tag{4.5}$$

Because both boundary equations, equations 4.4 and 4.5, are homogenous in G_{Na}, G_{K}, and G_{Leak}, we can describe the boundary plane with two variables. For example, we can rewrite equation 4.4 as

$$\begin{aligned} &A_1(V^*)G_{\mathrm{Na}} + B_1(V^*)G_{\mathrm{K}} + G_{\mathrm{Leak}} = 0 \\ &A_1(V^*)\frac{G_{\mathrm{Na}}}{G_{\mathrm{Leak}}} + B_1(V^*)\frac{G_{\mathrm{K}}}{G_{\mathrm{Leak}}} = -1, \\ &A_1(V^*)N + B_1(V^*)K = -1, \end{aligned} \tag{4.6}$$

where N and K are conductance ratios. From equation 4.5, we obtain another constraining equation and must then solve the following system,

$$\begin{pmatrix} A_1(V^*) & B_1(V^*) \\ A_2(V^*) & B_2(V^*) \end{pmatrix} \begin{pmatrix} N \\ K \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \tag{4.7}$$

where A_1 and B_1 are coefficients of equation 4.4 as written explicitly in equation 4.6,

$$\begin{aligned} A_1(V) &= [m_\infty^3(V)]' h_\infty(V)(V - E_{\mathrm{Na}}) + m_\infty^3(V)h_\infty(V), \\ B_1(V) &= n_\infty^4(V), \end{aligned}$$

and A_2 and B_2 are coefficients of equation 4.5,

$$\begin{aligned} A_2(V) = {} & [m_\infty^3(V)]'' h_\infty(V)(V - E_{\mathrm{Na}}) + [m_\infty^3(V)]' [h_\infty(V)]'(V - E_{\mathrm{Na}}) \\ & + 2[m_\infty^3(V)]' h_\infty(V) + m_\infty^3(V)[h_\infty(V)]', \\ B_2(V) = {} & [n_\infty^4(V)]'. \end{aligned}$$

The general solution is:

$$\begin{aligned} N &= \frac{-B_2(V^*)}{A_1(V^*)B_2(V^*) - A_2(V^*)B_1(V^*)} \\ K &= \frac{A_2(V^*)}{A_1(V^*)B_2(V^*) - A_2(V^*)B_1(V^*)} \end{aligned}. \tag{4.8}$$

The forms of N and K are highly nonlinear due to the dependence on the unknown value V^* (see Figure 6a), where $V^* > -51$ mV for positive conductance ratios. Rather than solve for V^* for every parameter set, we observe that the relevant range of V^* is highly constrained by the need to generate conductance values in a physiological range. The range used,

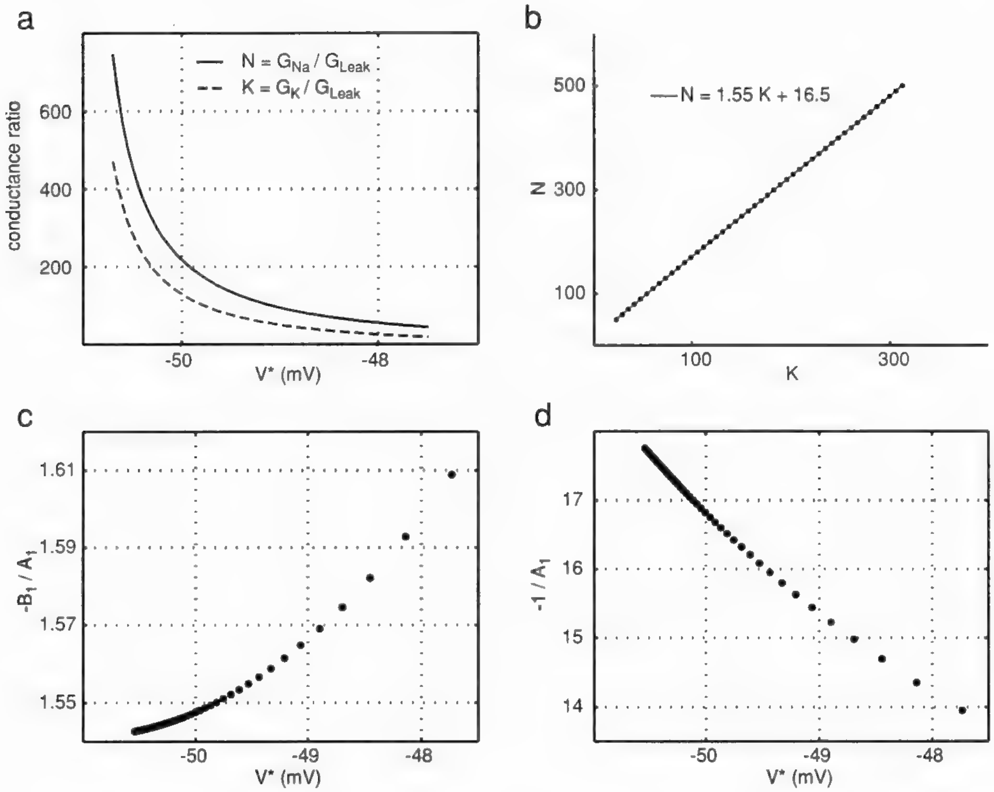


Figure 6: Conductance ratios $N = G_{\mathrm{Na}}/G_{\mathrm{Leak}}$ and $K = G_{\mathrm{K}}/G_{\mathrm{Leak}}$ plotted (a) as functions of V and (b) against each other. The coefficients of the linear regression fit are close to the simulated coefficients of equation 3.4. The voltage range plotted in *b* is from −50.5 to −48 mV, where lower voltages correspond to higher ratios. Coefficients (c) $-B_1/A_1$ and (d) $-1/A_1$ of equation 4.4 can be plotted as a function of V^*, where these values of V^* give ratios of $N = [50\ 500]$. For a given V^*, these coefficients can be used to find the coefficients of the boundary, equation 3.4. The values of V^* used in *c* and *d* are the same as those used in *b* to plot N and K, demonstrating that exponentially spaced V^* give rise to uniformly spaced values in (K, N).

$V^* = [-50.5\ -48]$, gives $N = [50\ 500]$, which was a constraint for the simulation. In this very narrow range of V^*, N and K are linearly related to one another; these are plotted against one another in Figure 6b. When we fit the line of N versus K, we obtain coefficients of approximately 1.55 and 16.5, which are very close to the coefficients of 1.54 and 16.7 from G_{K} and G_{Leak} in equation 3.4, respectively. Exact agreement between simulated and calculated results is probably prevented by imprecision in gathering the simulated boundary points.

The coefficients of G_{K} and G_{Leak} in equation 3.4 can be determined analytically from the coefficients A_1 and B_1 of equation 4.6. Specifically, for a given value V^*, the coefficient for G_{K} is $-B_1(V^*)/A_1(V^*)$ and the coefficient for G_{Leak} is $-1/A_1(V^*)$, as plotted in Figures 6c and 6d. Given equation 4.8, $-B_1/A_1$ and $-1/A_1$ are the slope and y-intercept of the linear relationship N versus K. Regression, as in Figure 6b, takes into account many values of V^* and does not require knowing V^* a priori, while using A_1 and B_1 gives an analytic result for a given V^*.

5 Hodgkin-Huxley Boundary Equation

The coefficients of 2.07 and 22.8 from G_K and G_{Leak} of the HH boundary equation, equation 3.1, differ by a constant from those (1.54 and 16.7) of the AK boundary, equation 3.4. The HH model is less excitable than the AK model, since less outward conductance is needed to suppress firing. In the approximations made in the 2D reduction of HH to AK, the largest source of error probably stems from the approximation of infinitely fast dynamics for the sodium activation variable, such that $m \approx m_\infty$. Around the typical voltage threshold, m has a time constant of approximately 0.3 to 0.5 msec. This will cause the true m to lag behind the approximated $m \approx m_\infty$ and to take a lower value at any point in time leading up to the threshold, causing the lower excitability of the HH model. Comparing m dynamics during spiking near the boundary with those of m_∞ (not shown), we find that multiplying m_∞ by a constant factor of three-fourths provides a better approximation to m around the threshold. When this factor is included in the AK model, the boundary equation coefficients extracted from simulations are found to be 2.09 and 22.14, in good agreement with those determined by the HH equations in equation 3.1. With this correction, the V-nullcline of equation 4.4 can be used to find the coefficients as in Figure 6b: 2.07 and 21.3.[5]

6 Slow Sodium Inactivation

While we have considered only steady-state conductances, this approach may help to conceptualize the behavior of a system subject to certain types of adaptation, such as slow sodium inactivation (Arsiero et al., 2007). That the addition of slow sodium inactivation effectively lowers maximal sodium conductance for steady state can be seen from the model equations. Slow sodium inactivation may be introduced (Fleidervish, Friedman, & Gutnick, 1996; Miles, Dai, & Brownstone, 2005) through an additional slow sodium inactivation variable s:

---

[5] This is, of course, not a true V-nullcline of the HH model since we are not specifying a dependent variable to plot it against. But as in the AK case, we assume that that the slow gating variables n and h are not functions of V and can then be ignored when partial differentiating with respect to V. This is analogous to assuming some sort of mapping from the 3D HH (i.e., excluding m) to a 2D model, and again we see that coefficients of equation 4.4, with the 3/4 correction factor multiplying m_∞, are related to the coefficients of equation 3.1.

$$I_{\mathrm{Na}} = G_{\mathrm{Na}} m^3 h s (V - E_{\mathrm{Na}}), \tag{6.1}$$

where s has the same form as the inactivation variable h but with a much slower time constant. The addition of s lowers the effective sodium pool for spiking. At steady state, when s is approximately constant due to its slow kinetics relative to spiking, this addition is equivalent to lowering G_{Na}. Through the sigmoidal voltage dependence of s, changes in mean current lead to effective changes of G_{Na}; as mean current increases, s decreases, decreasing the effective sodium pool. In other words, in contrast to spike-dependent adaptation currents, the voltage dependence of sodium inactivation provides feedback regardless of spiking activity. Thus, slow sodium inactivation, which in some sense is similar to voltage-dependent adaptation currents (Benda & Herz, 2003), is one mechanism by which the functional role of a neuron could change from integrating to differentiating without changes in channel density.

7 Discussion

Models of a variety of neurons share the same form as the Hodgkin-Huxley (HH) neuron (Ermentrout, 1998; Shriki, Hansel, & Sompolinsky, 2003), and evidence suggests that single-compartment models can capture key properties of in vivo and in vitro neurons (Destexhe et al., 2001). Here, we show specifically how a ratio of inward and outward conductances affects the HH neuron's sensitivity to input variance. The balance of outward to inward currents affects neuronal excitability. The well-known coincidence detection of auditory neurons (Reyes, Rubel, & Spain, 1996; Trussell, 1997) is enhanced by both a low-threshold potassium channel, which effectively decreases the membrane time constant (Reyes et al., 1994; Rathouz & Trussell, 1998; Rothman & Manis, 2003), and low availability of sodium channels (Svirskis, Kotak, Sanes, & Rinzel, 2004). In dendrites, membrane excitability is modulated by changes in sodium or potassium currents (Colbert, Magee, Hoffman, & Johnston, 1997; Johnston, Hoffman, Colbert, & Magee, 1999). Low sodium conductance, which leads to spike frequency adaptation, may be the result of intrinsic low channel density (Melnick, Santos, & Safronov, 2004) or slow sodium inactivation (Fleidervish et al., 1996; Miles et al., 2005), which can be modulated by second messenger systems (Cantrell & Catterall, 2001).

Neurons demonstrate a spectrum of integration and differentiation, and slow adaptation currents may correlate with shifts along this spectrum, at least for more complicated neurons such as neocortical pyramidal neurons (Higgs et al., 2006). However, diverging *f-I* curves, as in Figure 2, would more likely be related to voltage-dependent rather than spike-dependent adaptation. Previous work has demonstrated that increased shunting, a likely consequence of strong excitatory and inhibitory synaptic input, can allow *M*-current adaptation to effectively switch a neuron's operating regime from integration to coincidence detection, or differentiation (Prescott et al., 2006). This may be a specific case of our general result, where the *M*-current essentially increases G_K and shunting increases G_{Leak}. In some cases, the synaptic conductance alone may be large enough to switch a neuron from integration to differentiation. For example, Destexhe and Paré (1999) found that the input resistance of pyramidal cells in the parietal cortex decreases approximately fivefold during intense network activity. In the standard HH model, our results indicate that this change would place the cell near the boundary between integrator and differentiator function. Because synaptic conductances are highly variable on both short and long timescales, a postsynaptic neuron may fluctuate dynamically across the functional boundary.

Although we divide neural behavior into integration and differentiation as classified by the variance dependence of *f-I* curves, there are other possible ways to determine the computation that a single neuron performs on its current inputs. One example is white noise analysis, which we have explored extensively in applications to the Hodgkin-Huxley model (Aguera y Arcas et al., 2003), reduced models (Aguera y Arcas & Fairhall, 2003; Hong et al., 2007), and neurons of avian brain stem (Slee, Higgs, Fairhall, & Spain, 2005). While we have found, suggestively, that the computation as described by a linear/nonlinear model changes as a function of G_{Na}/G_K (Mease, 2006), here we do not attempt to connect these results.

In conclusion, we find that a plane separates two computational regimes in the space of maximal ionic conductances in the Hodgkin-Huxley and Abbott-Kepler models. Hyperplanes have been noted to separate models with different firing characteristics (Goldman, Golowasch, Marder, & Abbott, 2001; Taylor, Hickey, Prinz, & Marder, 2006); in this case we show how the plane can be derived from characteristics of the *V* nullcline. Since *V* nullclines for reduced models of neuronal dynamics can in principle be obtained experimentally from $I(V)$ plots (Izhikevich, 2007), this method may provide a means of relating biophysical properties with computational ones. Specifically, one may be able to infer certain biophysical parameters and their related effect on the neuron's computation from an experimentally obtained set of $I(V)$ plots. Finally, although we focus here on maximal conductances, as might be regulated by homeostatic mechanisms, many forms of adaptation, such as slow sodium inactivation, change conductances over time, so that each adaptation state has a different maximal conductance. Time-dependent movement of the system through different computational states may be an effective way to describe the functional role of adaptation.

Appendix A: Biophysical Modeling

The single-compartmental conductance-based Hodgkin-Huxley (HH) model neuron (Hodgkin & Huxley, 1952) was used with standard parameters (Koch, 1999; Dayan & Abbott, 2001; Gerstner & Kistler, 2002) except as noted. In addition to equation 1.1, the following equations comprise the HH model:

$$\frac{dn}{dt} = \alpha_n(1-n) - \beta_n n,$$

$$\frac{dm}{dt} = \alpha_m(1-m) - \beta_m m,$$

$$\frac{dh}{dt} = \alpha_h(1-h) - \beta_h h,$$

$$\alpha_n(V) = \frac{0.01(V+55)}{1-e^{-0.1(V+55)}} \qquad \beta_n(V) = 0.125e^{-(V+65)/80},$$

$$\alpha_m(V) = \frac{0.1(V+40)}{1-e^{-0.1(V+40)}} \qquad \beta_m(V) = 4e^{-(V+65)/18},$$

$$\alpha_h(V) = 0.07e^{-(V+65)/20} \qquad \beta_h(V) = \frac{1}{1+e^{-0.1(V+35)}},$$

where steady-state gating values, such as n_∞, are equal to $\alpha/(\alpha+\beta)$. Standard parameters for HH are: $G_{\mathrm{Na}} = 120$, $G_{\mathrm{K}} = 36$, and $G_{\mathrm{Leak}} = 0.3$ mS/cm^2; $E_{\mathrm{Na}} = 50$, $E_{\mathrm{K}} = -77$, and $E_{\mathrm{Leak}} = -54.4$ mV; and $C = 1\,\mu$F/cm^2. For the cortical neuron model of Figure 1, standard parameters are: $G_{\mathrm{Na}} = 50$, $G_{\mathrm{Kfast}} = 225$, $G_{\mathrm{Kslow}} = 0.225$, and $G_{\mathrm{Leak}} = 0.25$ mS/cm^2; $E_{\mathrm{Na}} = 74$, $E_{\mathrm{Kfast}} = E_{\mathrm{Kslow}} = -90$, and $E_{\mathrm{Leak}} = -70$ mV; and $C = 1\,\mu$F/cm^2 with channel kinetics as in Table 4.1 (p. 124) of Gerstner and Kistler (2002).

Equations for the Abbott-Kepler model are derived from the HH model (Abbott & Kepler, 1990; Kepler et al., 1992; Hong et al., 2007). Equations were solved numerically using fourth-order Runge-Kutta integration with a fixed time step of 0.05 msec, or 0.025 msec for the AK model. Injected current was simulated by a series of normally distributed random numbers that were smoothed by an exponential filter ($\tau = 1$ msec). Spike times were identified as the upward crossing of the voltage trace at -20 mV (resting potential = -65 mV) separated by more than 2 msec.

Appendix B: Examining $f(U, V)$

That equation 4.4 is the appropriate condition can also be seen as follows. While the function f depends on both U and V, along the nullclines, U is a function of V, $U = g(V)$. Thus, requiring $df(U, V)/dV = 0$ at the fixed point as required by the definition of a V nullcline, we have:

$$\begin{aligned}\frac{df(U,V)}{dV} &= \frac{\partial f}{\partial V} + \frac{\partial f}{\partial U}\frac{dU}{dV} = 0,\\ \frac{dU}{dV} &= \frac{-\frac{\partial f}{\partial V}}{\frac{\partial f}{\partial U}} = 0,\end{aligned} \tag{B.1}$$

where $\partial f/\partial U > 0$. Then the numerator must equal zero, giving equation 4.4. One can show that the full stability analysis around a fixed point gives almost the same result as our discussion based on the shape of a nullcline. For the 2D AK model, the following two conditions on the determinant and trace of the Jacobian matrix J evaluated at the fixed point are met for a stable fixed point (Strogatz, 1994):

$$\begin{aligned}\det(J) &= \frac{\partial F}{\partial V}\frac{\partial g}{\partial U} - \frac{\partial F}{\partial U}\frac{\partial g}{\partial V} > 0,\\ \mathrm{Tr}(J) &= \frac{\partial F}{\partial V} + \frac{\partial g}{\partial U} < 0,\end{aligned} \tag{B.2}$$

where $F(U,V) = -f(U,V)/C + I/C = dV/dt$, given equation 4.2, and $g(U,V) = dU/dt$. The first condition is always satisfied for AK models, as the Jacobian at fixed points of these models always have complex eigenvalues due to their resonating dynamics. For the U nullcline of the AK model, $g \approx -k\,(U - V)$, where k has the form:

$$k = \frac{G_K F_K(V) t_K(V) + G_{Na} F_{Na}(V) t_{Na}(V)}{G_K F_K(V) + G_{Na} F_{Na}(V)}.$$

The functions $t_i(V)$ are slowly varying functions that are of the same order of magnitude over the range of V that we are interested in. Therefore, k weakly depends on conductances and is roughly constant in V. Moreover, the value of k is small with a maximum value less than 0.3 msec^{-1}. Therefore, $\frac{\partial f}{\partial V} = -C\frac{\partial F}{\partial V} \geq -0.3\,\text{mS/cm}^2$ is approximately equivalent to a full stability analysis based on the Jacobian. Even when k is nonzero, there is not a qualitative change; since $\frac{\partial f}{\partial V} = G_{Leak} + (V$ dependent terms), considering nonzero constant k is equivalent to $G_{Leak} \rightarrow G_{Leak} + Ck$, which corresponds to a small offset of our boundary plane such that it does not pass through the origin. Thus, the right-hand side of equations 3.1 and 3.4 may not equal zero. When we fit our simulated data and allow for a nonzero offset, the coefficients and error of the fit are almost identical as previously, and the fitted offset is small (HH: ∼−0.3; AK: ∼1.0).

Acknowledgments

We thank W. J. Spain, Rebecca Mease, Michele Giugliano, and Larry Sorenson for helpful discussions and Michele Giugliano for comments on the manuscript. This work was supported by a Burroughs-Wellcome Careers at the Scientific Interface grant and a Sloan Research Fellowship to A.L.F.; B.N.L. was supported by the Medical Scientist Training Program, a fellowship from the National Institute of General Medical Sciences (T32 07266), and an ARCS fellowship; M.H.H. was supported by a VA Merit Review to W.J.S. This project was also supported by grant number F30NS055650 (to B.N.L.) from the National Institute of Neurological Disorders and Stroke. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Neurological Disorders and Stroke or the National Institutes of Health.

References

Abbott, L. F., & Kepler, T. B. (1990). Model neurons: From Hodgkin-Huxley to Hopfield. In: L. Garrido (Ed.), *Statistical mechanics of neural networks* (pp. 5–18). Berlin: Springer-Verlag.

Abeles, M. (1982). Role of the cortical neuron: Integrator or coincidence detector? *Isr. J. Med. Sci., 18*, 83–92.

Aguera y Arcas, B., & Fairhall, A. L. (2003). What causes a neuron to spike? *Neural Comput., 15*, 1789–1807.

Aguera y Arcas, B., Fairhall, A. L., & Bialek, W. (2003). Computation in a single neuron: Hodgkin and Huxley revisited. *Neural Comput., 15*, 1715–1749.

Arsiero, M., Luscher, H. R., Lundstrom, B. N., & Giugliano, M. (2007). The impact of input fluctuations on the frequency-current relationships of layer 5 pyramidal neurons in the rat medial prefrontal cortex. *J. Neurosci., 27*, 3274–3284.

Benda, J., & Herz, A. V. (2003). A universal model for spike-frequency adaptation. *Neural Comput., 15*, 2523–2564.

Bryant, H. L., & Segundo, J. P. (1976). Spike initiation by transmembrane current: A white-noise analysis. *J. Physiol., 260*, 279–314.

Cantrell, A. R., & Catterall, W. A. (2001). Neuromodulation of Na+ channels: An unexpected form of cellular plasticity. *Nat. Rev. Neurosci., 2*, 397–407.

Chance, F. S., Abbott, L. F., & Reyes, A. D. (2002). Gain modulation from background synaptic input. *Neuron, 35*, 773–782.

Colbert, C. M., Magee, J. C., Hoffman, D. A., & Johnston, D. (1997). Slow recovery from inactivation of Na+ channels underlies the activity-dependent attenuation of dendritic action potentials in hippocampal CA1 pyramidal neurons. *J. Neurosci., 17*, 6512–6521.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press.

Destexhe, A., & Paré, D. (1999). Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo. *J. Neurophysiol., 81*, 1531–1547.

Destexhe, A., Rudolph, M., Fellous, J. M., & Sejnowski, T. J. (2001). Fluctuating synaptic conductances recreate in vivo–like activity in neocortical neurons. *Neuroscience*, *107*, 13–24.

Destexhe, A., Rudolph, M., & Paré, D. (2003). The high-conductance state of neocortical neurons in vivo. *Nat. Rev. Neurosci.*, *4*, 739–751.

Ermentrout, B. (1998). Linearization of F-I curves by adaptation. *Neural Comput.*, *10*, 1721–1729.

Fellous, J. M., Rudolph, M., Destexhe, A., & Sejnowski, T. J. (2003). Synaptic background noise controls the input/output characteristics of single cells in an in vitro model of in vivo activity. *Neuroscience*, *122*, 811–829.

Fleidervish, I. A., Friedman, A., & Gutnick, M. J. (1996). Slow inactivation of Na+ current and slow cumulative spike adaptation in mouse and guinea-pig neocortical neurones in slices. *J. Physiol.*, *493*(Pt. 1), 83–97.

Gerstein, G. L., & Mandelbrot, B. (1964). Random walk models for the spike activity of a single neuron. *Biophys. J.*, *4*, 41–68.

Gerstner, W., & Kistler, W. M. (2002). *Spiking neuron models: Single neurons, populations, plasticity.* Cambridge: Cambridge University Press.

Giugliano, M., Bove, M., & Grattarola, M. (1999). Activity-driven computational strategies of a dynamically regulated integrate-and-fire model neuron. *J. Comput. Neurosci.*, *7*, 247–254.

Goldman, M. S., Golowasch, J., Marder, E., & Abbott, L. F. (2001). Global structure, robustness, and modulation of neuronal models. *J. Neurosci.*, *21*, 5229–5238.

Gonsalves, B., & Paller, K. A. (2000). Neural events that underlie remembering something that never happened. *Nat. Neurosci.*, *3*, 1316–1321.

Gutkin, B., Ermentrout, G. B., & Rudolph, M. (2003). Spike generating dynamics and the conditions for spike-time precision in cortical neurons. *J. Comput. Neurosci.*, *15*, 91–103.

Higgs, M. H., Slee, S. J., & Spain, W. J. (2006). Diversity of gain modulation by noise in neocortical neurons: Regulation by the slow afterhyperpolarization conductance. *J. Neurosci.*, *26*, 8787–8799.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, *117*, 500–544.

Hong, S., Aguera y Arcas, B., & Fairhall, A. L. (2007). Single neuron computation: From dynamical system to feature detector. *Neural Comput.*, *19*, 3133–3172.

Izhikevich, E. M. (2007). *Dynamical systems in neuroscience: The geometry of excitability and bursting.* Cambridge, MA: MIT Press.

Johnston, D., Hoffman, D. A., Colbert, C. M., & Magee, J. C. (1999). Regulation of back-propagating action potentials in hippocampal neurons. *Curr. Opin. Neurobiol.*, *9*, 288–292.

Kepler, T. B., Abbott, L. F., & Marder, E. (1992). Reduction of conductance-based neuron models. *Biol. Cybern.*, *66*, 381–387.

Koch, C. (1999). *Biophysics of computation: Information processing in single neurons.* New York: Oxford University Press.

Konig, P., Engel, A. K., & Singer, W. (1996). Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends Neurosci.*, *19*, 130–137.

LeMasson, G., Marder, E., & Abbott, L. F. (1993). Activity-dependent regulation of conductances in model neurons. *Science, 259,* 1915–1917.

Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science, 268,* 1503–1506.

Mease, R. (2006). [Simulation results showing that the linear/nonlinear model obtained for Hodgkin-Huxley models with variations in the maximal sodium and potassium conductances are invariant along lines of constant conductance ratios.] Unpublished data.

Melnick, I. V., Santos, S. F., & Safronov, B. V. (2004). Mechanism of spike frequency adaptation in substantia gelatinosa neurones of rat. *J. Physiol., 559,* 383–395.

Miles, G. B., Dai, Y., & Brownstone, R. M. (2005). Mechanisms underlying the early phase of spike frequency adaptation in mouse spinal motoneurones. *J. Physiol., 566,* 519–532.

Moreno, R., de la Rocha, J., Renart, A., & Parga, N. (2002). Response of spiking neurons to correlated inputs. *Phys. Rev. Lett., 89,* 288101.

Murray, J. D. (2002). *Mathematical biology* (3rd ed.). New York: Springer.

Oertel, D. (1983). Synaptic responses and electrical properties of cells in brain slices of the mouse anteroventral cochlear nucleus. *J. Neurosci., 3,* 2043–2053.

Prescott, S. A., Ratte, S., De Koninck, Y., & Sejnowski, T. J. (2006). Nonlinear interaction between shunting and adaptation controls a switch between integration and coincidence detection in pyramidal neurons. *J. Neurosci., 26,* 9084–9097.

Rathouz, M., & Trussell, L. (1998). Characterization of outward currents in neurons of the avian nucleus magnocellularis. *J. Neurophysiol., 80,* 2824–2835.

Rauch, A., La Camera, G., Luscher, H. R., Senn, W., & Fusi, S. (2003). Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo–like input currents. *J. Neurophysiol., 90,* 1598–1612.

Reyes, A. D., Rubel, E. W., & Spain, W. J. (1994). Membrane properties underlying the firing of neurons in the avian cochlear nucleus. *J. Neurosci., 14,* 5352–5364.

Reyes, A. D., Rubel, E. W., & Spain, W. J. (1996). In vitro analysis of optimal stimuli for phase-locking and time-delayed modulation of firing in avian nucleus laminaris neurons. *J. Neurosci., 16,* 993–1007.

Richardson, M. J. (2004). Effects of synaptic conductance on the voltage distribution and firing rate of spiking neurons. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys., 69,* 051918.

Richardson, M. J., & Gerstner, W. (2005). Synaptic shot noise and conductance fluctuations affect the membrane voltage with equal significance. *Neural. Comput., 17,* 923–947.

Rothman, J. S., & Manis, P. B. (2003). The roles potassium currents play in regulating the electrical activity of ventral cochlear nucleus neurons. *J. Neurophysiol., 89,* 3097–3113.

Roy, S. A., & Alloway, K. D. (2001). Coincidence detection or temporal integration? What the neurons in somatosensory cortex are doing. *J. Neurosci., 21,* 2462–2473.

Rudolph, M., & Destexhe, A. (2003a). Characterization of subthreshold voltage fluctuations in neuronal membranes. *Neural Comput., 15,* 2577–2618.

Rudolph, M., & Destexhe, A. (2003b). Tuning neocortical pyramidal neurons between integrators and coincidence detectors. *J. Comput. Neurosci., 14,* 239–251.

Shriki, O., Hansel, D., & Sompolinsky, H. (2003). Rate models for conductance-based cortical neuronal networks. *Neural Comput.*, *15*, 1809–1841.

Slee, S. J., Higgs, M. H., Fairhall, A. L., & Spain, W. J. (2005). Two-dimensional time coding in the auditory brainstem. *J. Neurosci.*, *25*, 9978–9988.

Strogatz, S. H. (1994). *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. Reading, MA: Addison-Wesley.

Svirskis, G., Kotak, V., Sanes, D. H., & Rinzel, J. (2004). Sodium along with low-threshold potassium currents enhance coincidence detection of subthreshold noisy signals in MSO neurons. *J. Neurophysiol.*, *91*, 2465–2473.

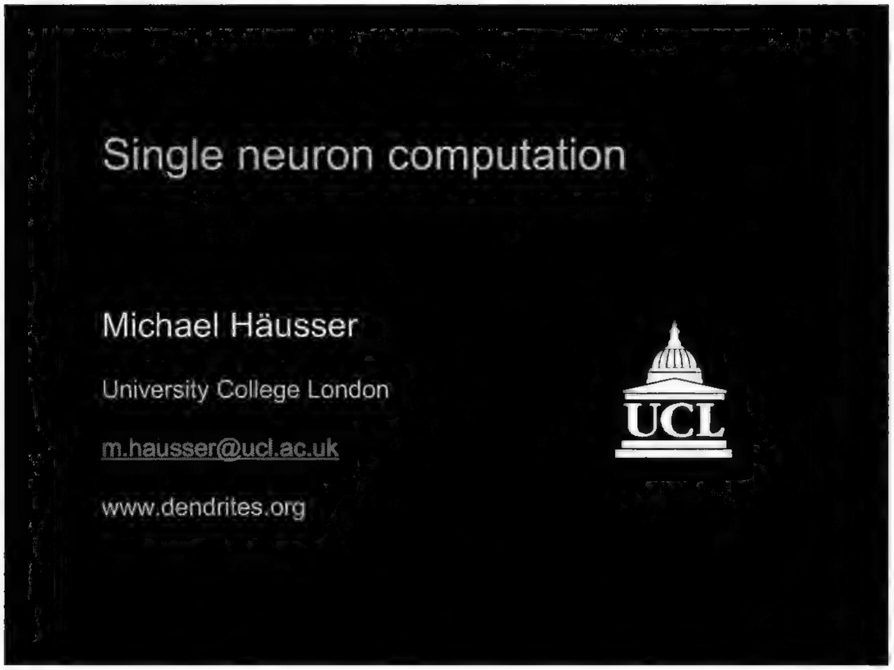
Taylor, A. L., Hickey, T. J., Prinz, A. A., & Marder, E. (2006). Structure and visualization of high-dimensional conductance spaces. *J. Neurophysiol.*, *96*, 891–905.

Trussell, L. O. (1997). Cellular mechanisms for preservation of timing in central auditory pathways. *Curr. Opin. Neurobiol.*, *7*, 487–492.

Tuckwell, H. C. (1988). *Introduction to theoretical neurobiology.* Cambridge: Cambridge University Press.

Yu, Y., & Lee, T. S. (2003). Dynamical mechanisms underlying contrast gain control in single neurons. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.*, *68*, 011901.

Received May 29, 2007; accepted July 16, 2007.



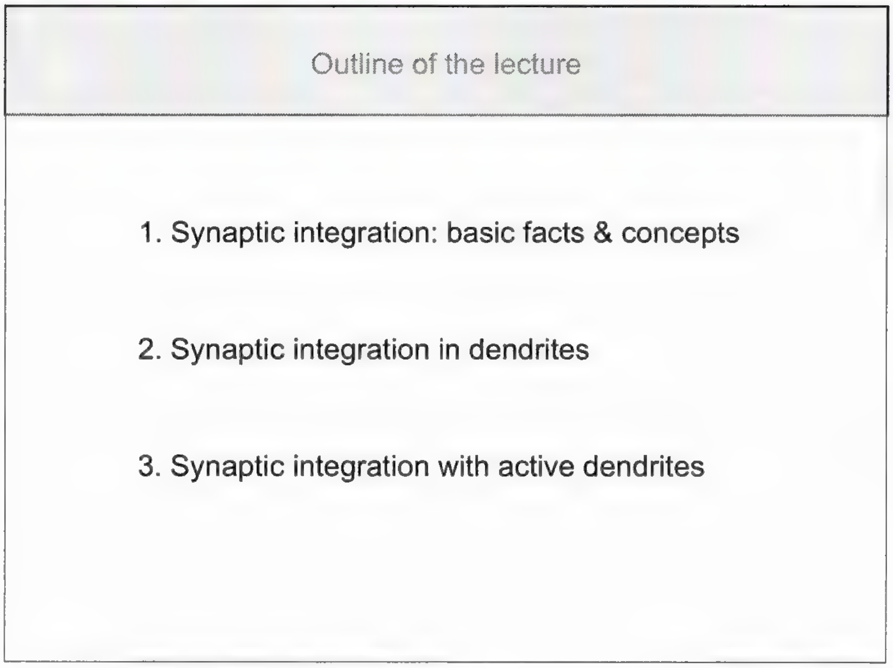
Single neuron computation

Michael Häusser

University College London

m.hausser@ucl.ac.uk

www.dendrites.org



Outline of the lecture

1. Synaptic integration: basic facts & concepts
2. Synaptic integration in dendrites
3. Synaptic integration with active dendrites

Figure 2.2
Camillo Golgi (1843–1926). [illegible] Finger, 1994, Fig. [illegible]

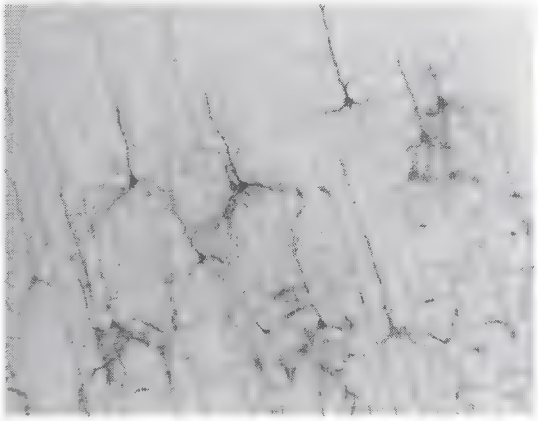
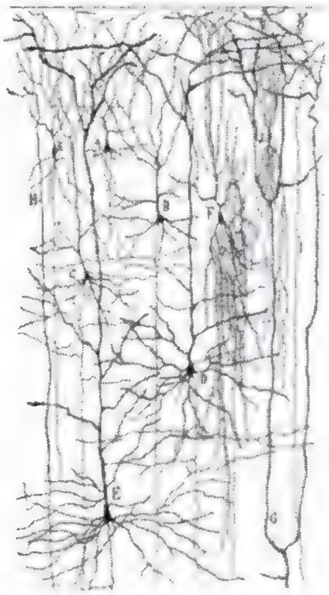
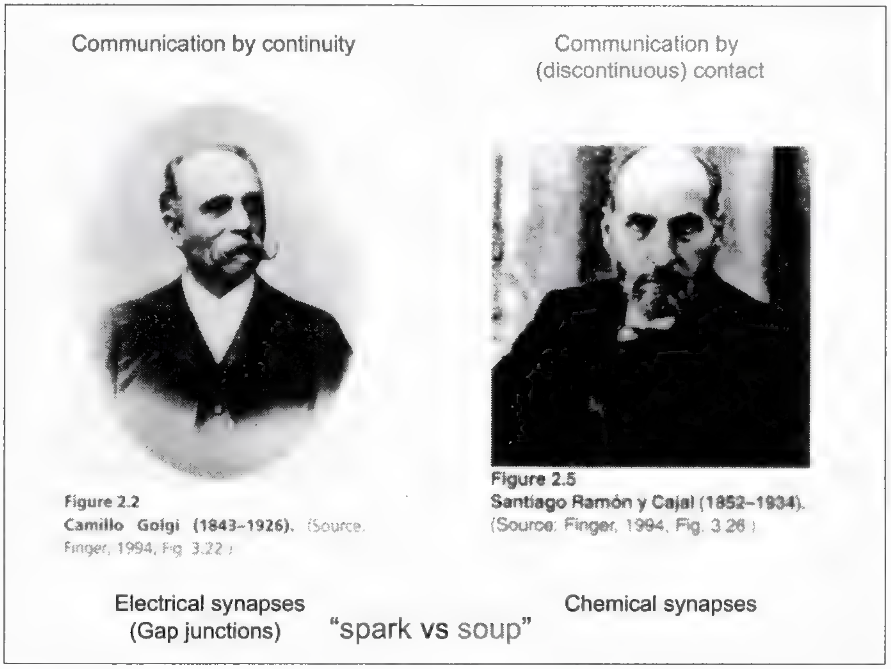


Figure 2.3
Golgi-stained neurons. [illegible]

Figure 2.5
Santiago Ramon y Cajal (1852–1934).
(Source: Finger, 1994, Fig. 3.26.)





Communication by continuity

Communication by (discontinuous) contact

Figure 2.2
Camillo Golgi (1843–1926). (Source. Finger, 1994, Fig 3.22)

Figure 2.5
Santiago Ramón y Cajal (1852–1934). (Source: Finger, 1994, Fig. 3.26)

Electrical synapses
(Gap junctions)

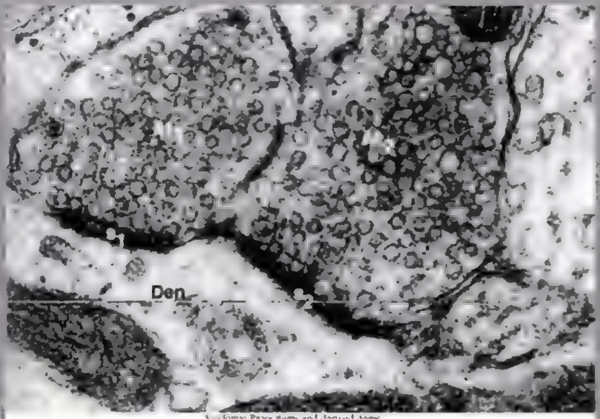
“spark vs soup”

Chemical synapses

Neuron doctrine

The nervous system is divided into independent functional entities, *neurons*.

Neurons are the fundamental anatomical, physiological, metabolic, and genetic units of the nervous system.

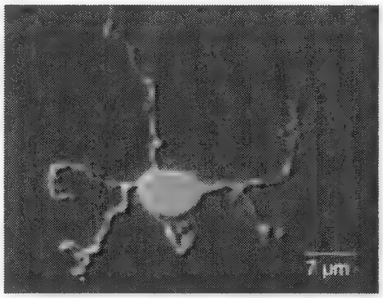


Academic Press items and derived items
copyright © 1999 by Academic Press

The first electron micrographs (1950s) confirmed Cajal's neuron doctrine.

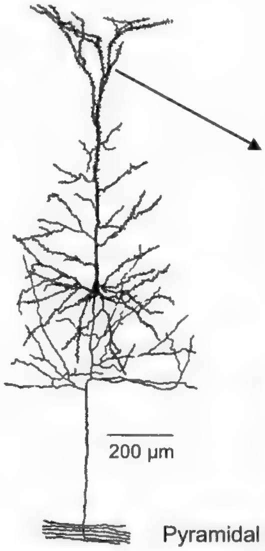
How complicated should single neurons be?

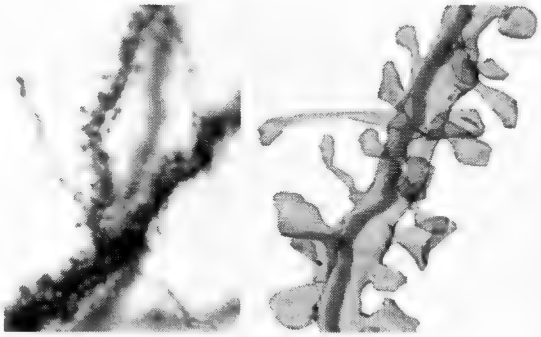
Some neurons are simple…..



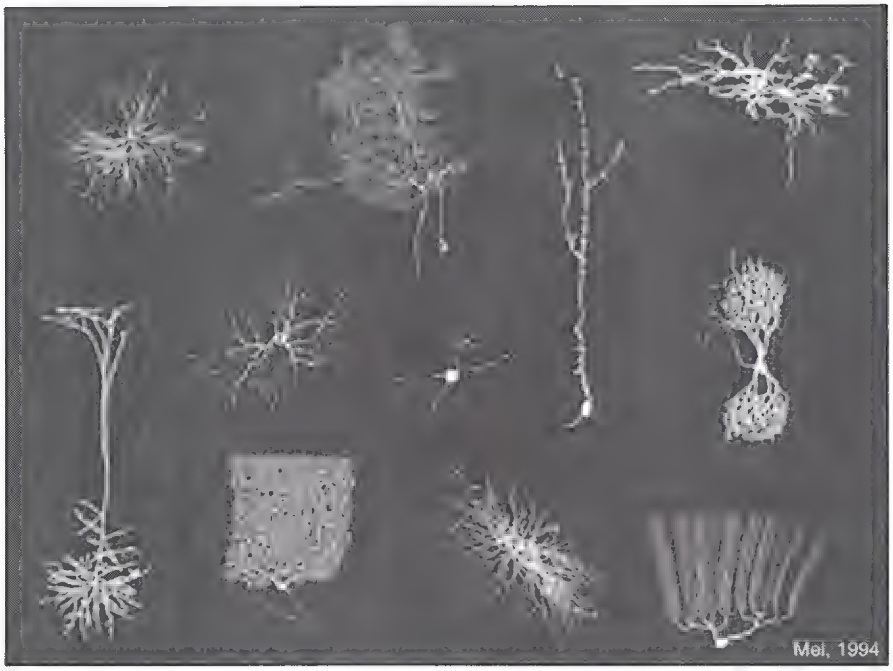
Granule cell, cerebellar cortex

…while others are highly complex.





Pyramidal cell, cerebral cortex

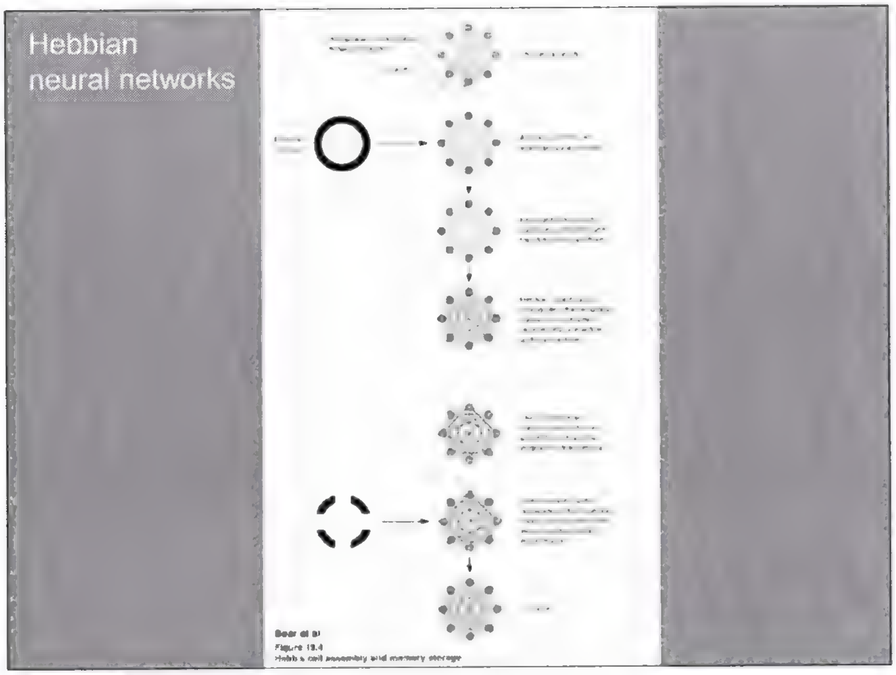


Mel, 1994

[illegible]

Two strategies:

1. Simple, "dumb" units; information processing primarily carried out by the network of connections (e.g. cerebellar granule cells)



Hebbian neural networks

Bear et al

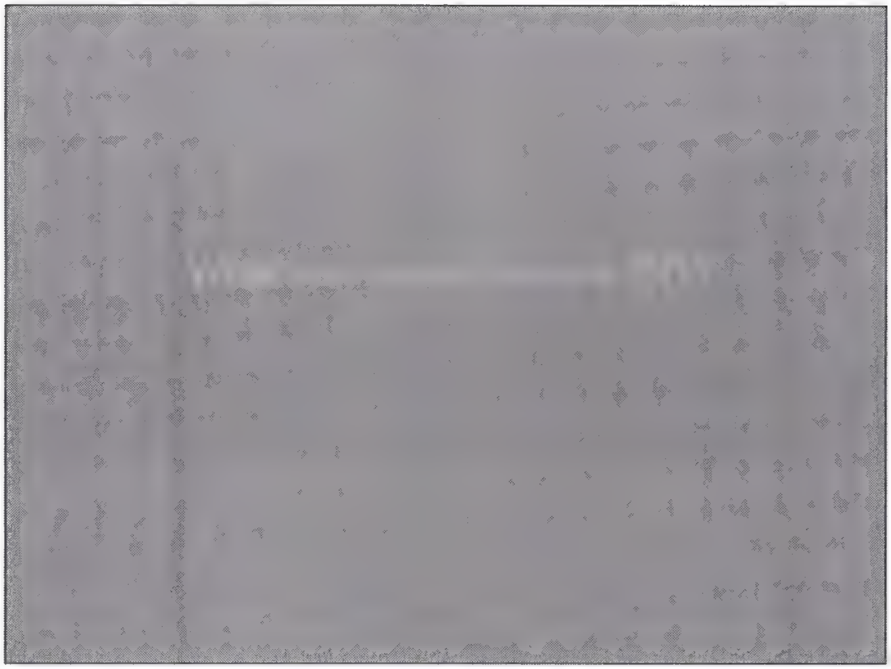
Figure 19.4

Hebb's cell assembly and memory storage

[illegible]

Two strategies:

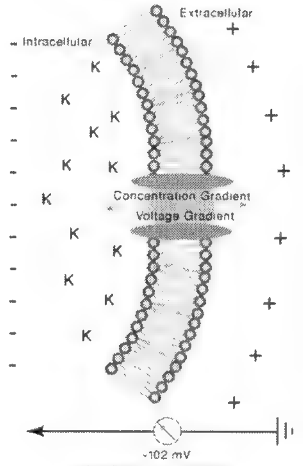
1. Simple, "dumb" units; information processing primarily carried out by the network of connections (e.g. cerebellar granule cells).
2. Complex, "smart" units, capable of sophisticated computation on their inputs (e.g. cortical pyramidal cells).



Basic facts:

1. Most neurons have a resting membrane potential of around **-65 mV**.

Basis of the membrane potential



Academic Press items and derived items
copyright © 1999 by Academic Press

- **"Electricity is only charge (a mystery), the separation of charge (voltage) and the movement of charge (current)."**
  - Louis de Felice, *Electrical Properties of Cells.*
- Charge is separated across the cell membrane, producing a transmembrane voltage (= the membrane potential).
- Many ions have a different concentration inside and outside the cell, producing a concentration gradient.
- Consider K+. [K] in >> [K] out.
- When K channels open, K+ can flow down its concentration gradient, leaving the cell.
  - This is the movement of charge, i.e., a K current
- This changes the charge separation across the membrane.
  - The voltage changes.
  - Note: usually, the number of ions that cross the membrane is so small relative to the total number of ions, that the concentration does not appreciably change.

The Nernst Equation

$$E_{ion} = \frac{RT}{zF} \bullet \ln\left(\frac{[ion]_o}{[ion]_i}\right)$$

E_{ion} is in Volts

R = gas constant = 8.315 J/K·mol

T = temperature in Kelvin
= T(°C) + 273.16

Z = valence of ion of interest

F = Faraday's constant
= 9.648 e4 C/mol

Important ions in neurons

	[out]	[in]	E
	(mM)	(mM)	(mV)
K^+	5	100	-80

The Goldman-Hodgkin-Katz (GHK) Equation

$$V_m = \frac{RT}{F} \bullet \ln\left(\frac{P_K[K^+]_o + P_{Na}[Na^+]_o + P_{Cl}[Cl^-]_i}{P_K[K^+]_i + P_{Na}[Na^+]_i + P_{Cl}[Cl^-]_o}\right)$$

$$= 61.54mV \bullet \log\left(\frac{P_K[K^+]_o + P_{Na}[Na^+]_o + P_{Cl}[Cl^-]_i}{P_K[K^+]_i + P_{Na}[Na^+]_i + P_{Cl}[Cl^-]_o}\right)$$

$$= 61.54mV \bullet \log\left(\frac{[K^+]_o + b[Na^+]_o + c[Cl^-]_i}{[K^+]_i + b[Na^+]_i + c[Cl^-]_o}\right)$$

where $b = P_{Na}/P_K$ and $c = P_{Cl}/P_K$

Ions and the resting potential

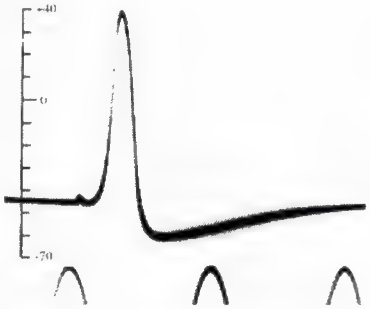
	[out]	[in]	E	P/PK
	(mM)	(mM)	(mV)	
K^+	5	100	-80	1
Na^+	150	15	+62	0.025
Ca^{2+}	2	0.0002	+123	–
Cl^-	150	13	-65	0.1

V_m = -65 mV

Basic facts:

1. Most neurons have a resting membrane potential of around **-65 mV**.
2. Most neurons have an action potential threshold of around **-45 mV**.

The most famous action potential of all



Hodgkin & Huxley, *Nature* 1939

Overton (1902): Na^+ necessary for nerve and muscle to fire action potentials

Hodgkin & Huxley (1939), Curtis & Cole (1940): action potential overshoots 0 mV, therefore more than just a breakdown of V_m - approaches E_{Na}, suggesting an increased P_{Na}

Hodgkin & Katz (1949): amplitude of action potential is affected by $[Na^+]_o$ (predicted by GHK equation)

Some basic facts:

1. Typical resting membrane potential: **-65 mV**.
2. Typical action potential threshold: **-45 mV**.
3. As a consequence, **most neurons are silent**, and **require synaptic input to generate output.**

Summation of synaptic inputs is required to generate output

Neuromuscular Junction (NMJ)

- large numbers of boutons per axon
- large number of release sites per bouton
- large quantal size (> 1 mV)

⇒ each presynaptic AP produces a large EPSP
⇒ reliable link between presynaptic & postsynaptic APs

Central synapses

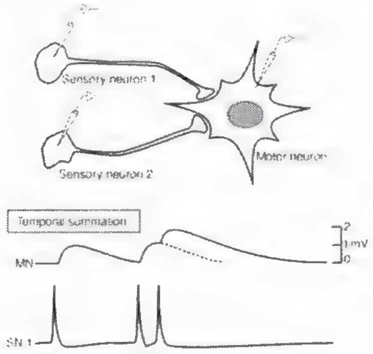
- only 1-5 boutons per axon on each postsynaptic cell
- only one release site per bouton
- small quantal size (~100 μV EPSP)

⇒ each presynaptic AP produces a small EPSP
⇒ single presynaptic inputs can't trigger postsynaptic APs

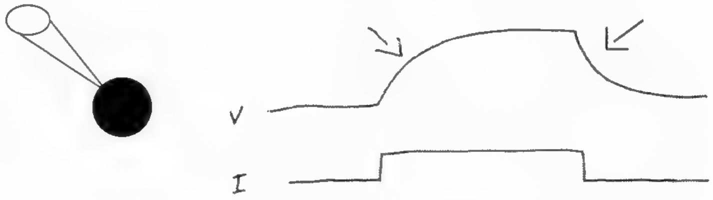
Why is this interesting?

⇒ because neurons can perform **computations** on their inputs, via their **combination** and **timing**

Temporal and Spatial Summation


Sensory neuron 1
Sensory neuron 2
Motor neuron
Temporal summation
MN
SN 1
2
1 mV
0

What determines the time window for summation?



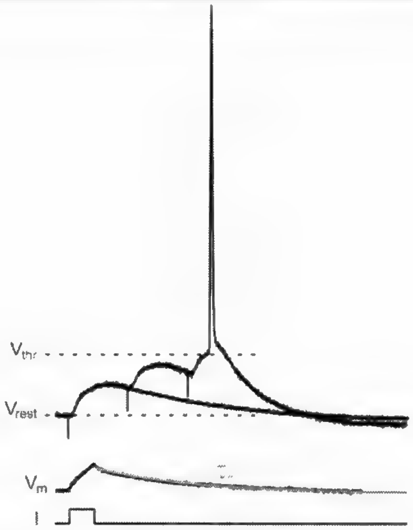
$$V = V_0 \cdot (1 - e^{-t/\tau})$$

$$\tau_m = R_m C_m \quad = 10 - 150 \text{ ms}$$

variable: 10,000-150,000 Ωcm^2

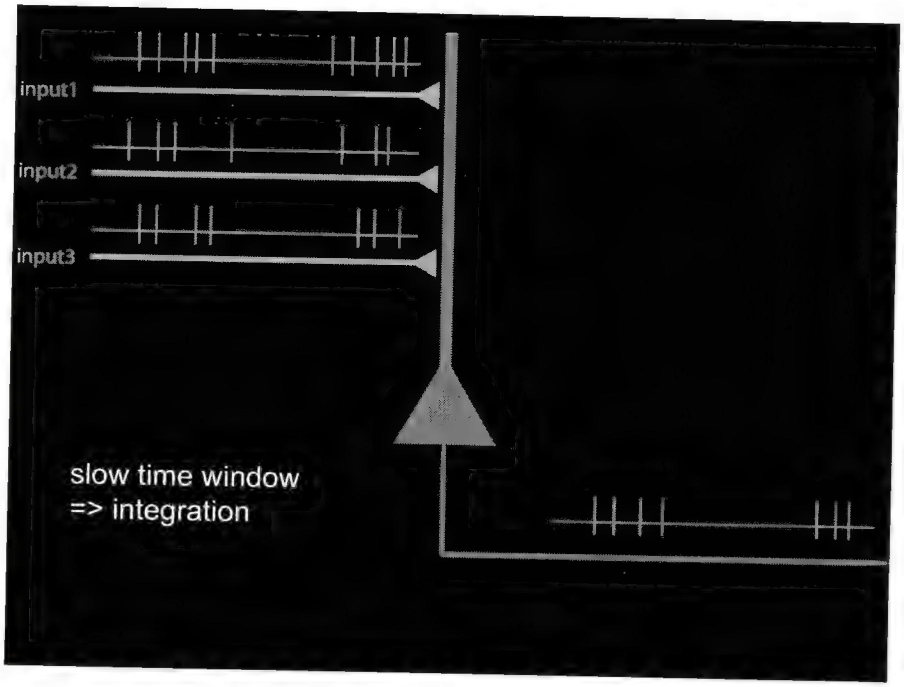
constant: 1 $\mu F/cm^2$

The membrane time constant can set the time window of summation

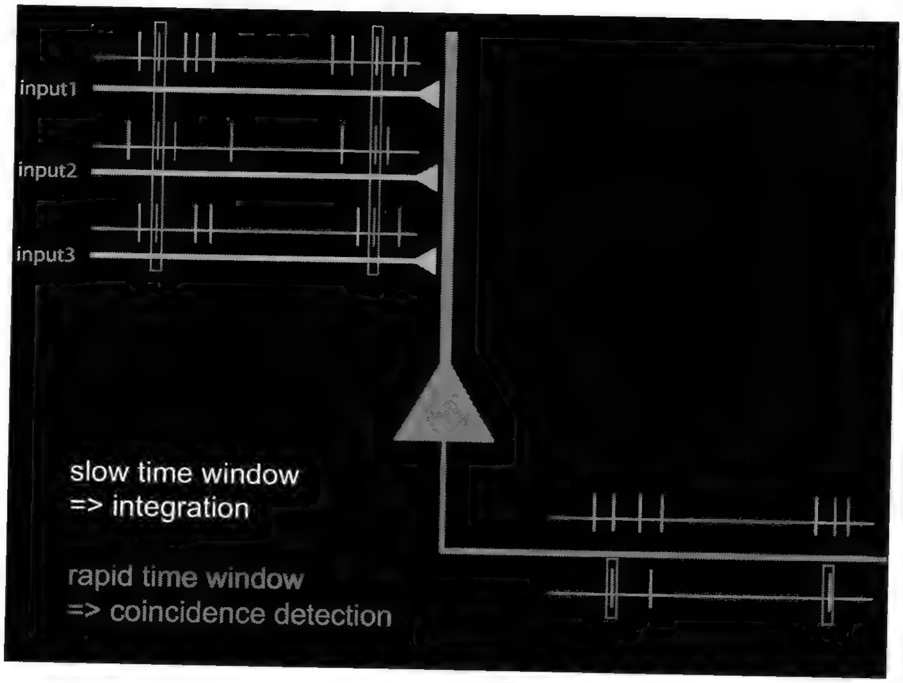


M. Scanziani

Why is the time window for summation important?


input1
input2
input3

slow time window
=> integration


input1
input2
input3

slow time window
=> integration

rapid time window
=> coincidence detection

Synaptic inhibition works in two ways

1. Hyperpolarizing inhibition

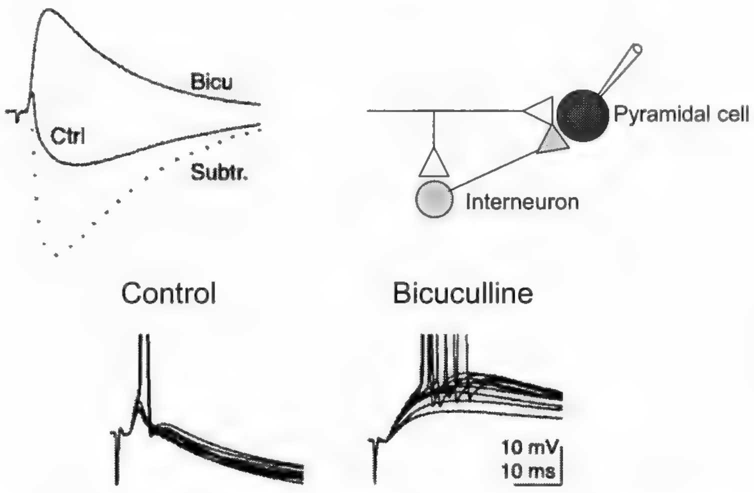
   If the reversal potential for the channels is hyperpolarized to the membrane potential

2. Shunting inhibition

   If the reversal potential for the channels is equal to the membrane potential

   - brief (only while synaptic conductance is on)
   - spatially localized
   - “divisive” rather than subtractive

Precisely timed inhibition can shorten the time window for integration



Pouille & Scanziani, 2001

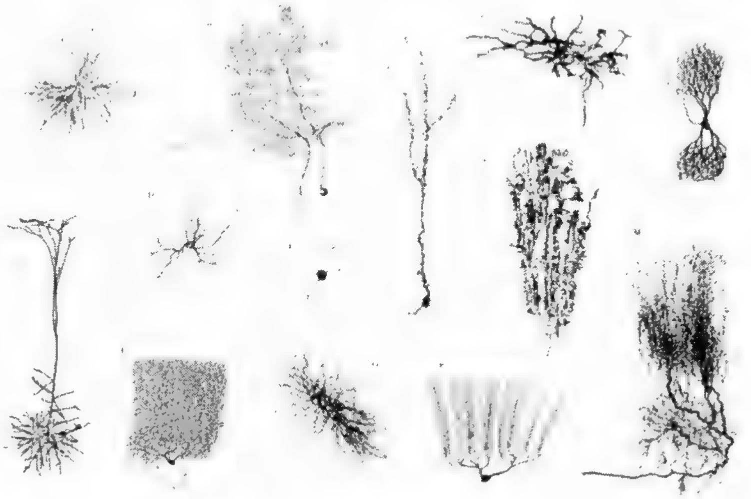
Outline of the lecture

1. Synaptic integration: basic facts & concepts

2. Synaptic integration in dendrites

3. Synaptic integration with active dendrites

Dendrites come in a variety of shapes and sizes



Current flow in neurons with dendrites

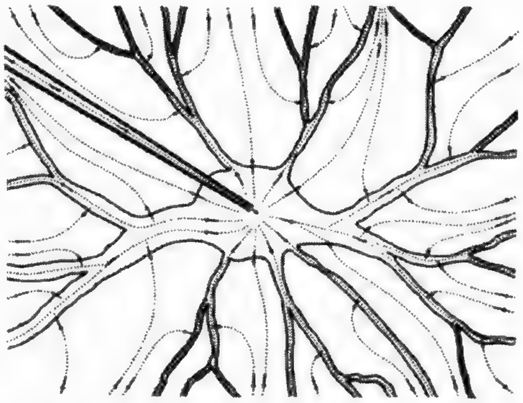
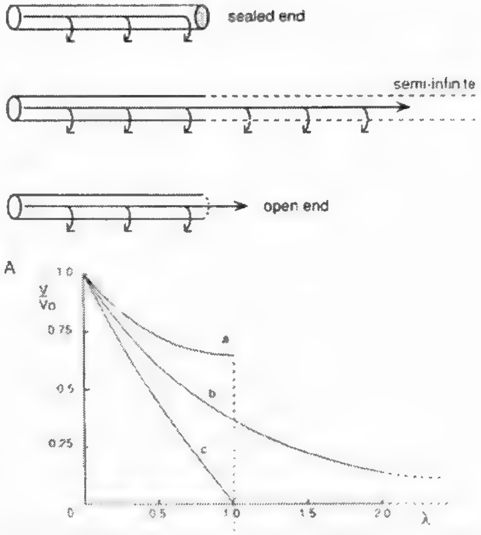


FIG. 1. Diagram illustrating the flow of electric current from a microelectrode whose tip penetrates the cell body (soma) of a neuron. The full extent of the dendrites is not shown. The external electrode to which the current flows is at a distance far beyond the limits of this diagram.

Wilfrid Rall

Voltage attenuation in different classes of cables



space constant

$$V = V_0 \cdot (1 - e^{-x/\lambda})$$

$$\lambda = \sqrt{\frac{R_m \cdot d}{R_i \cdot 4}}$$

Voltage attenuation in branching cables

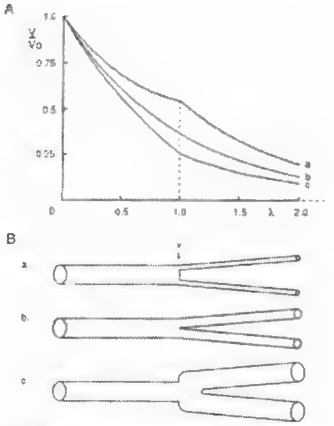
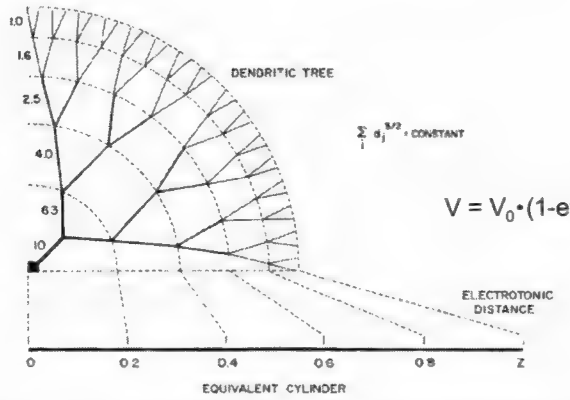


FIG. 13.7 Effect of different modes of dendritic branching on the spread of electrotonic potential. A, Graph of steady-state potential spread for the three cases illustrated in B. B, Diagrams illustrating three basic modes of branching; in each case the stem diameter is 4 μm. (a) Each daughter branch is 1 μm in diameter; (b) each daughter branch is approximately 2.5 μm, so that sum of $d^{3/2}$ equals $d^{3/2}$ of the stem; (c) each daughter branch has the same diameter as the stem. (Modified from Rall, 1959.)

$$d_1^{3/2} = d_{21}^{3/2} + d_{22}^{3/2}$$

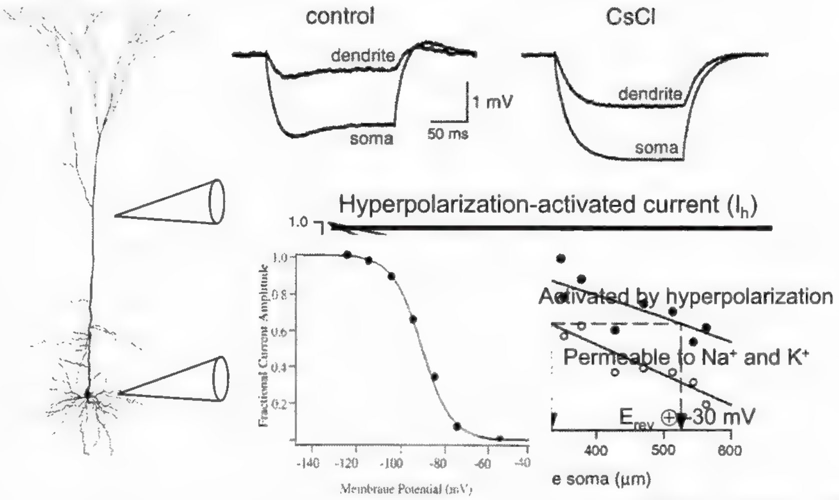
The equivalent cylinder model



$$V = V_0 \cdot (1 - e^{-x/\lambda})$$

FIG. 4. A dendritic tree illustrative of a class of trees that can be mathematically transformed into an equivalent cylinder. The diagram corresponds to a specific numerical example (Rall, 1962a, Table I) of a symmetric tree with a radial extent of about 800 μ, and with cylindrical diameters decreasing from a trunk of 10 μ to peripheral branches of 1 μ. The dashed lines connect points having the same Z-value (electrotonic distance) in both tree and cylinder. Z is defined by

Direct measurement of steady-state voltage attenuation in layer 5 pyramidal neurons



Hyperpolarization-activated current (I_h)

Activated by hyperpolarization

Permeable to Na^+ and K^+

$E_{rev} \approx -30$ mV

Stuart & Spruston, 1998

Compartmental modeling of neurons

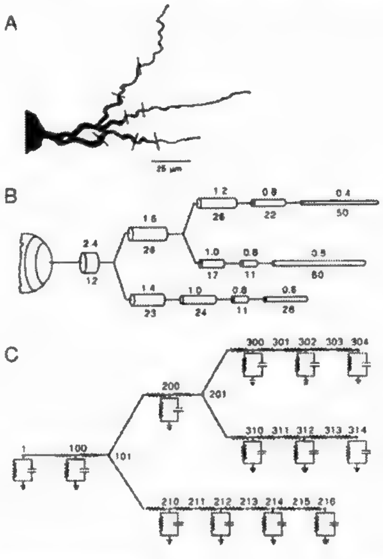
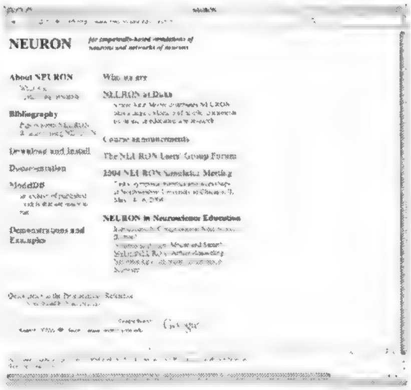

A
25 μm
B
2.4
12
1.6
26
1.2
26
0.8
22
0.4
50
1.0
17
0.8
11
0.5
60
1.4
23
1.0
24
0.8
11
0.6
26
C
1
100
101
200
201
300
301
302
303
304
310
311
312
313
314
210
211
212
213
214
215
216

Figure 3.3
Stages in abstraction from an anatomical dendritic tree to an electrical circuit analog. (A) Two-dimensional projection of part of the soma and one dendrite of a vagal motoneuron in the guinea pig. Points at which unbranched dendrites were broken into successive cylindrical segments are indicated by lines. (B) Representation of the same dendrite as a branched system of cylindrical segments, indicating the length (below) and diameter (above) of each dendritic segment (in μm). Diameters are not drawn to the same scale as the lengths, but both are in the correct proportions. The motoneuron soma (shown partially) has a maximum diameter of 20 μm and minimum diameter of 15 μm. (C) Circuit analog of B (see fig. 3.1) showing the pattern of connections at branch points and the numbers assigned to circuit nodes within (even numbers) and between (odd numbers) successive segments.

The NEURON simulation environment



NEURON

About NEURON

Bibliography

Download and Install

Documentation

ModelDB

Demonstrations and Examples

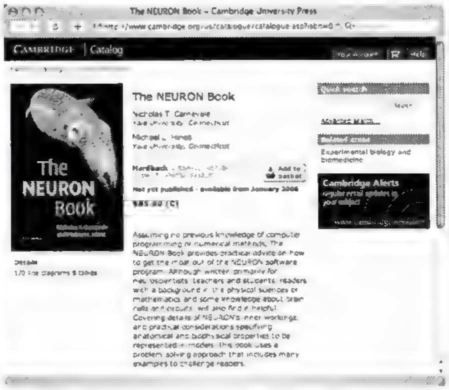
NEURON at Duke

Course announcements

The NEURON Users' Group Forum

2004 NEURON Simulator Meeting

NEURON in Neuroscience Education



The NEURON Book – Cambridge University Press

The NEURON Book

Nicholas T. Carnevale
Yale University, Connecticut

Michael L. Hines
Yale University, Connecticut

Hardback

Not yet published - available from January 2006

Experimental biology and biomedicine

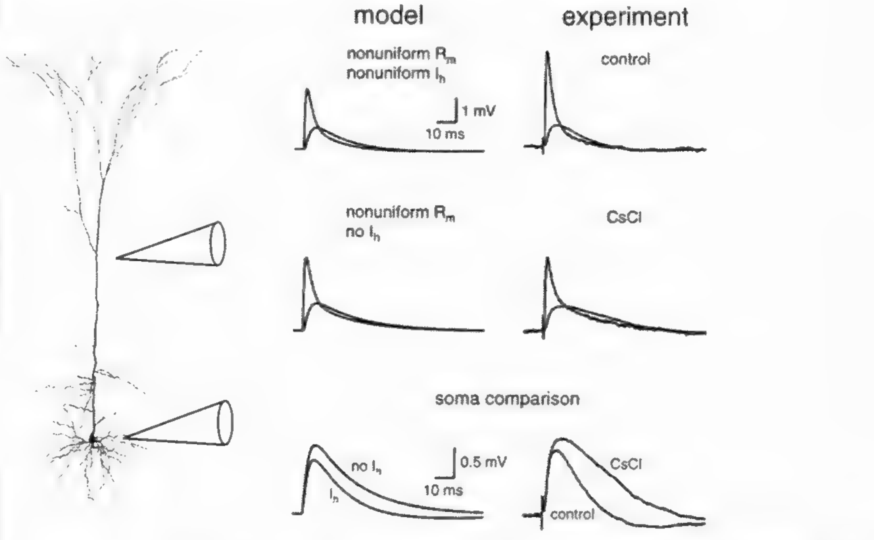
Cambridge Alerts
regular email updates in your subject

170 line diagrams 5 tables

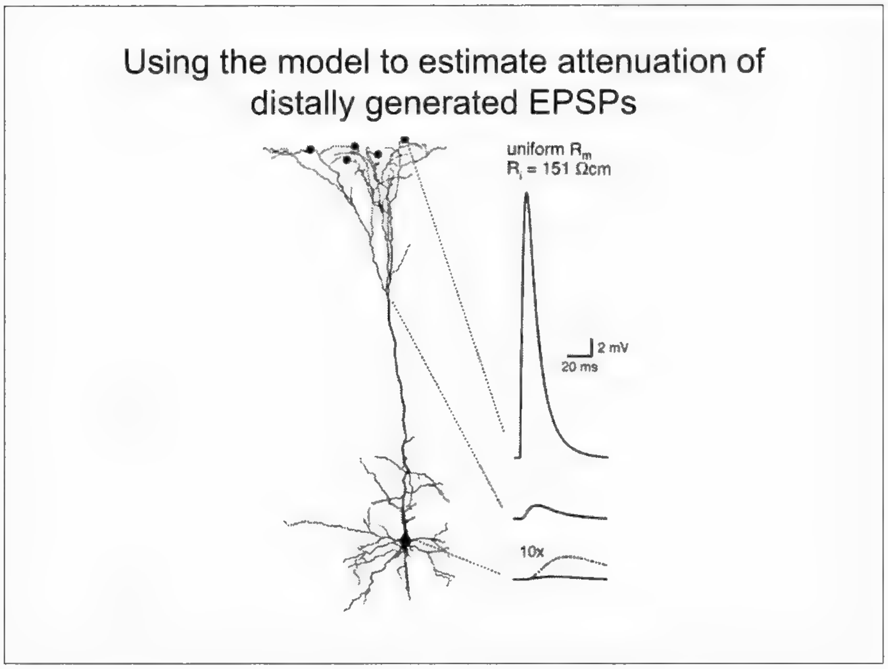
Assuming no previous knowledge of computer programming or numerical methods, The NEURON Book provides practical advice on how to get the most out of the NEURON software program. Although written primarily for neuroscientists, teachers and students, readers with a background in the physical sciences or mathematics and some knowledge about brain cells and circuits will also find it helpful. Covering details of NEURON's inner workings, and practical considerations specifying anatomical and biophysical properties to be represented in models, this book uses a problem solving approach that includes many examples to challenge readers.

www.neuron.yale.edu **Cambridge UP, 2006**

EPSP attenuation in dendrites of pyramidal neurons

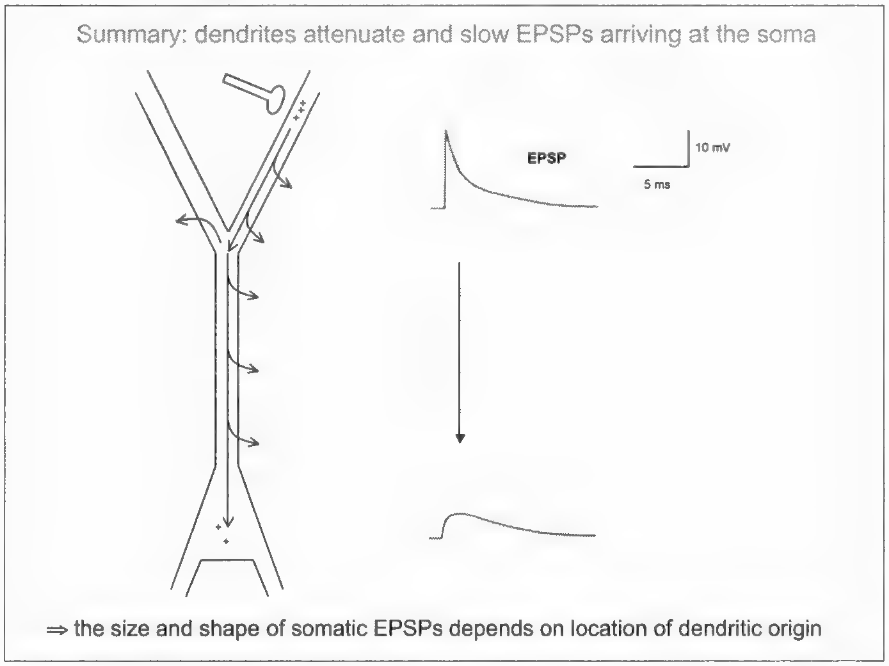


Stuart & Spruston, 1998



Using the model to estimate attenuation of distally generated EPSPs

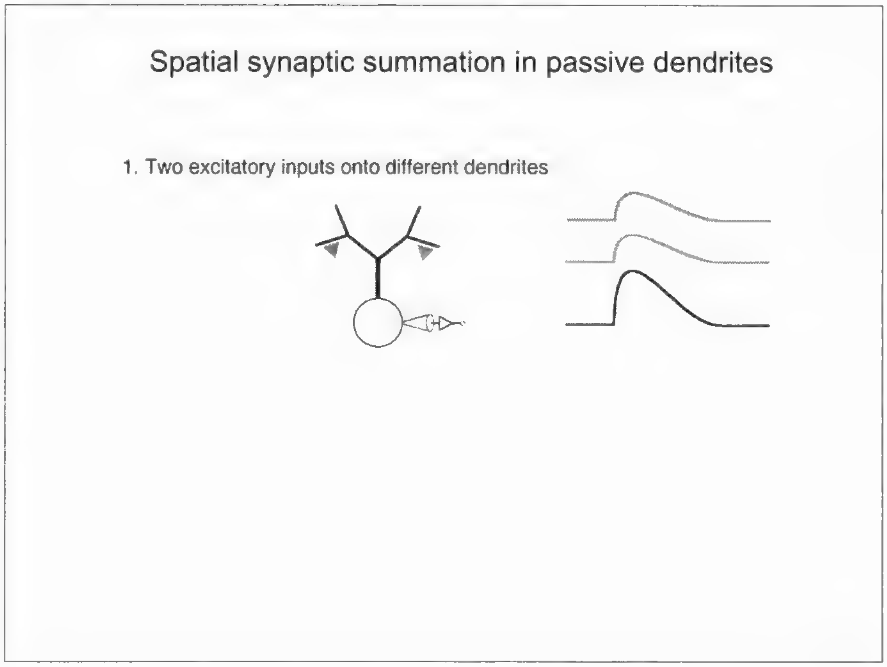
uniform R_m
R_i = 151 Ωcm
2 mV
20 ms
10x



Summary: dendrites attenuate and slow EPSPs arriving at the soma

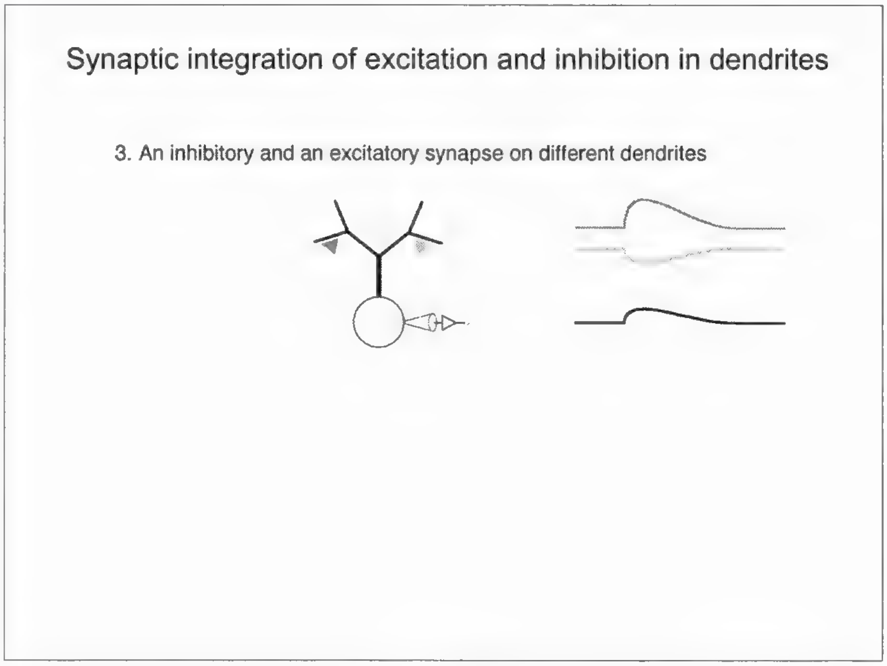
EPSP
10 mV
5 ms

⇒ the size and shape of somatic EPSPs depends on location of dendritic origin



Spatial synaptic summation in passive dendrites

1. Two excitatory inputs onto different dendrites



Synaptic integration of excitation and inhibition in dendrites

3. An inhibitory and an excitatory synapse on different dendrites

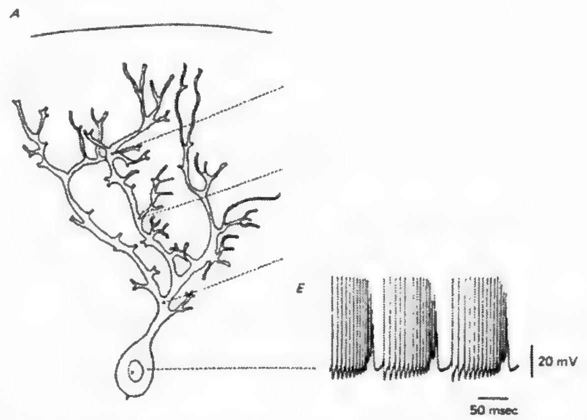
Outline of the lecture

1. Synaptic integration: basic facts & concepts

2. Synaptic integration in dendrites

3. Synaptic integration with active dendrites

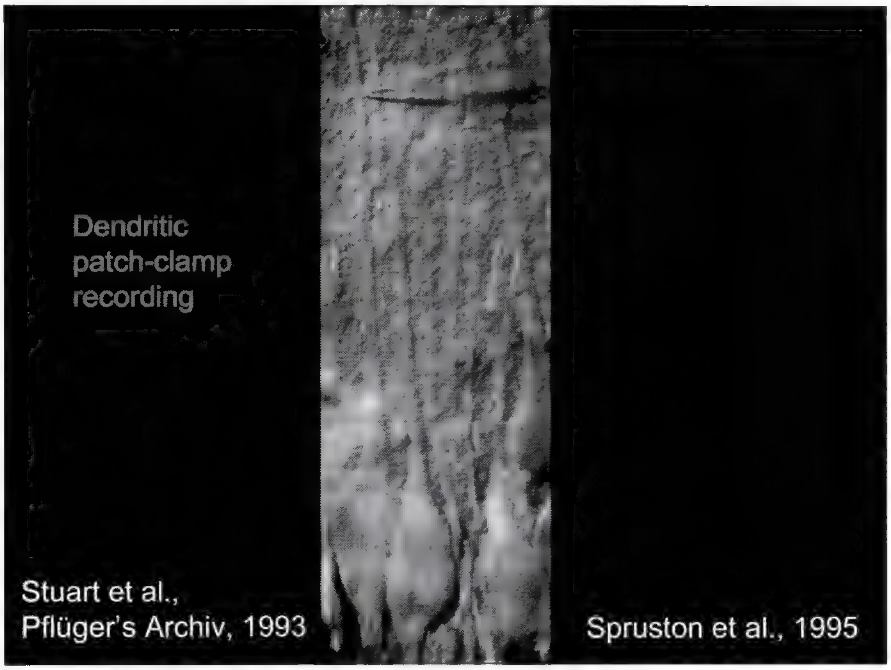
Compartmentalization of the dendritic tree by Na^+ and Ca^{2+} spikes



Na^+ spikes

Fig. 4. Composite picture showing the relationship between somatic and dendritic action potentials following DC depolarization through the recording electrode. A clear shift in amplitude of the s.s. against the dendritic Ca-dependent potentials is seen when comparing the more superficial recording in *B* with the somatic recording in *E*. Note that at increasing distances from the soma the fast spikes are reduced in amplitude and are barely noticeable in the more peripheral recordings. However, the prolonged and slow-rising burst spikes are more prominent at dendritic level.

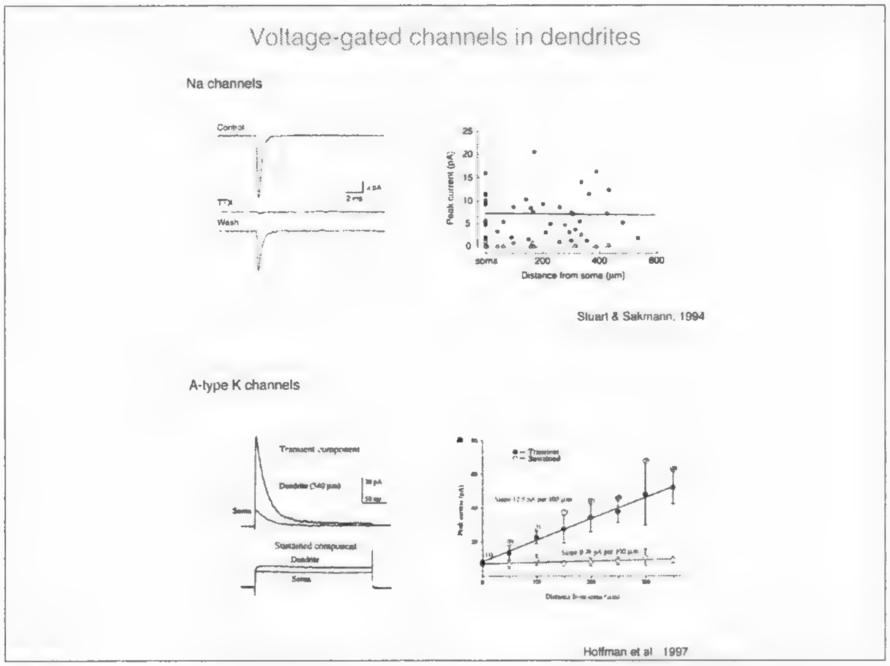
from Llinas & Sugimori 1980



Dendritic patch-clamp recording

Stuart et al.,
Pflüger's Archiv, 1993

Spruston et al., 1995



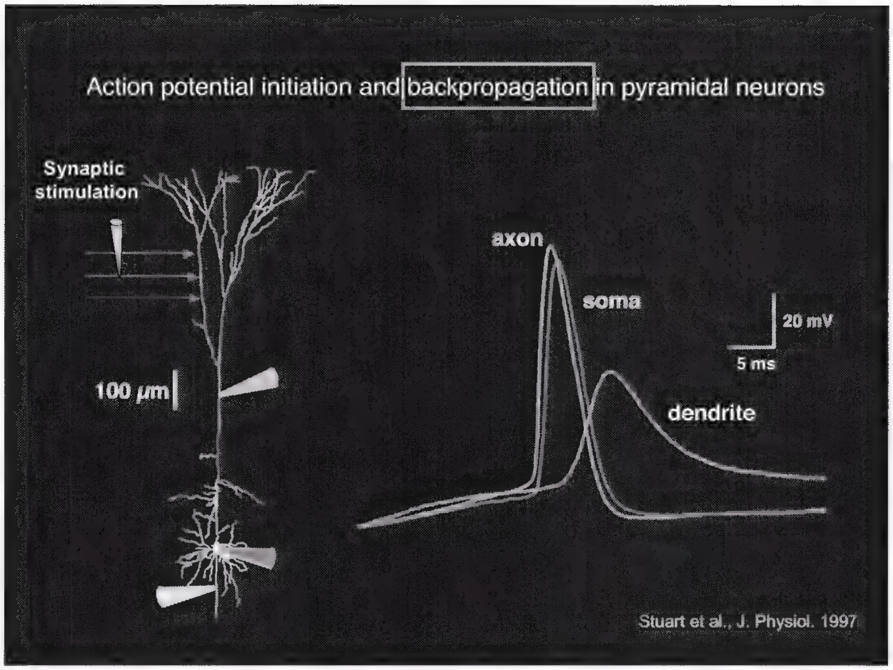
Voltage-gated channels in dendrites

Na channels

Stuart & Sakmann, 1994

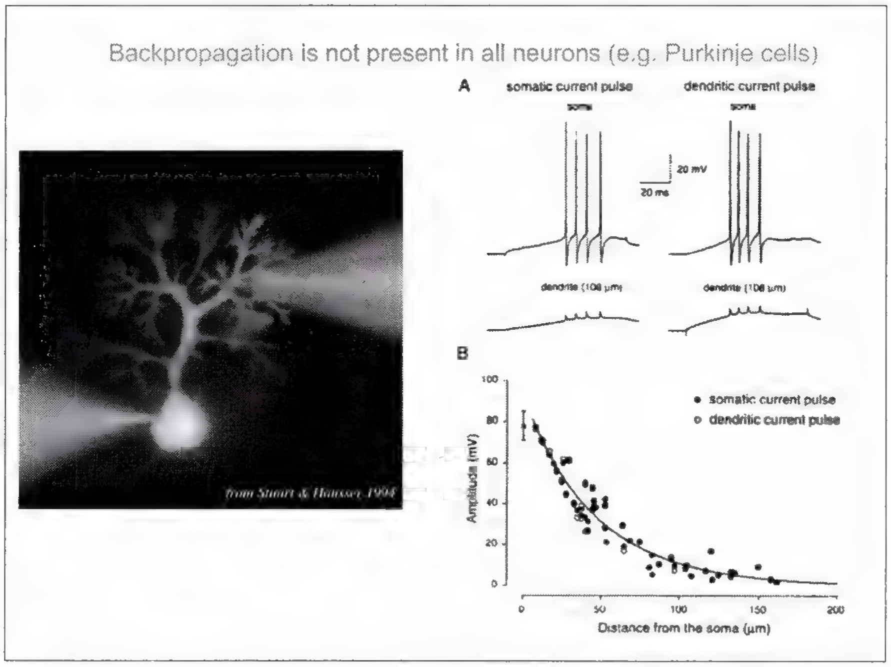
A-type K channels

Hoffman et al 1997



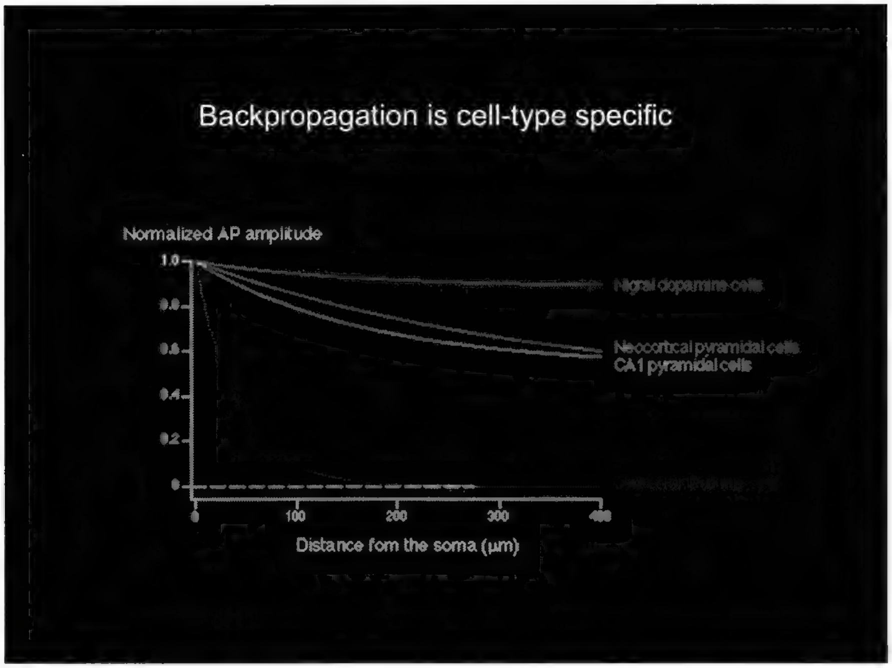
Action potential initiation and backpropagation in pyramidal neurons

Stuart et al., J. Physiol. 1997

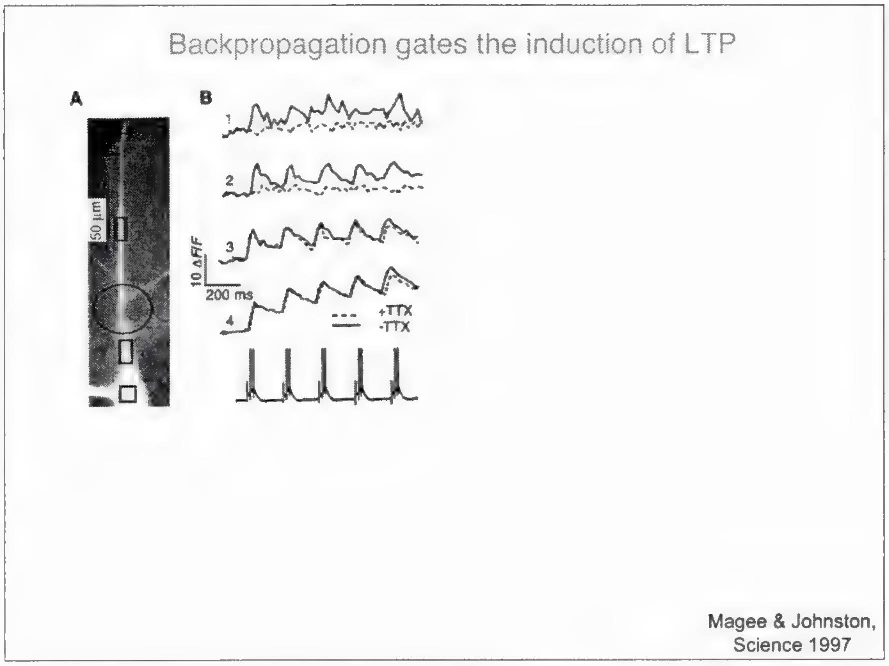


Backpropagation is not present in all neurons (e.g. Purkinje cells)

from Stuart & Häusser 1994



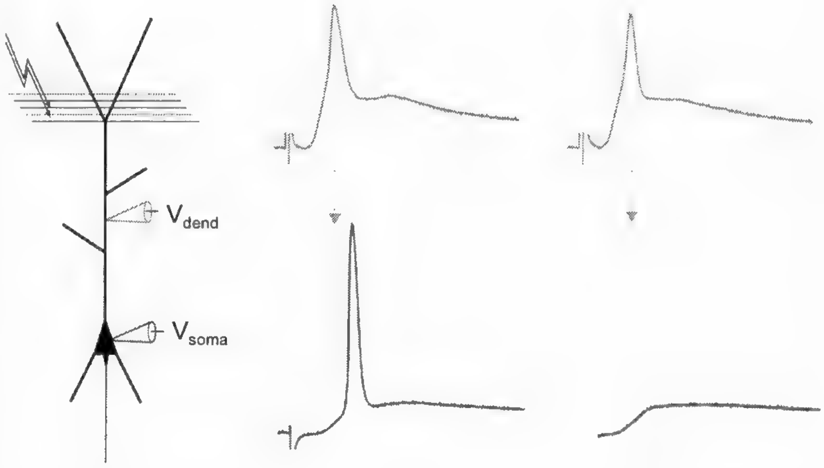
Backpropagation is cell-type specific



Backpropagation gates the induction of LTP

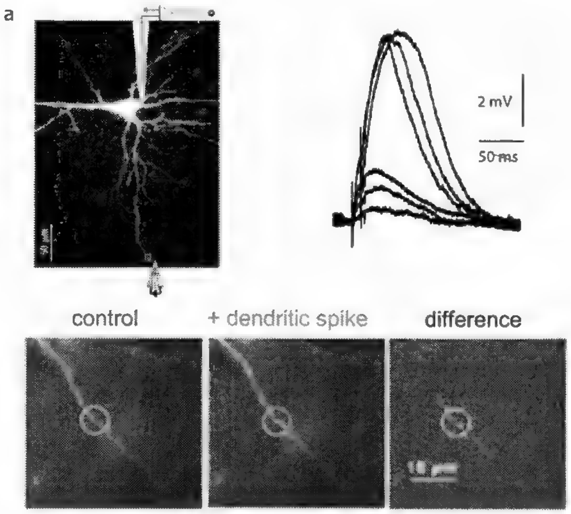
Magee & Johnston,
Science 1997

Dendrites can initiate local spikes which may or may not trigger output



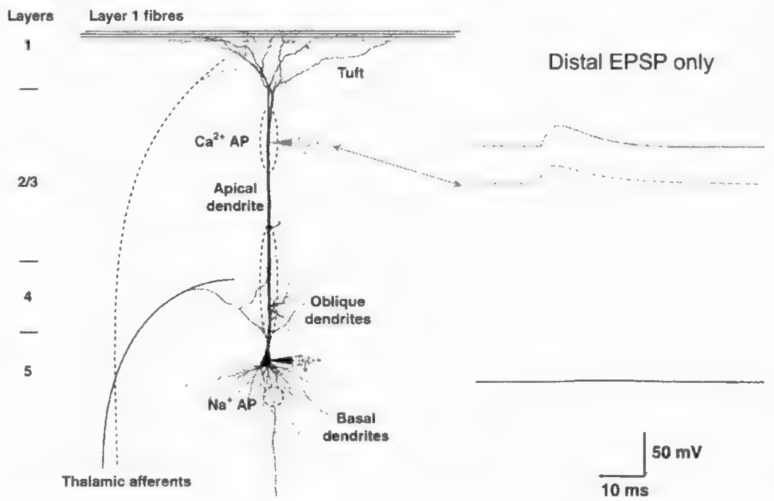
Golding and Spruston 1998

Dendritic spikes cause local calcium influx


a
50 µm
2 mV
50 ms
control
+ dendritic spike
difference
10 µm

Schiller et al., *Nature* 2000

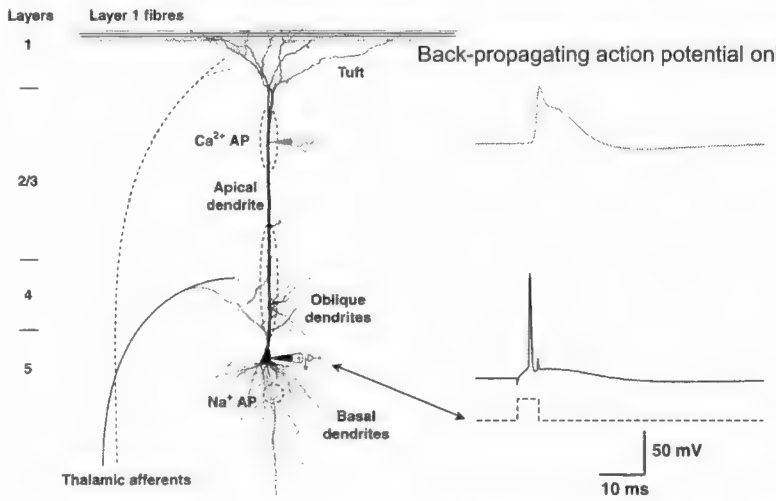
Interaction between backpropagating APs and distal Ca spikes


Distal EPSP only

Layers
1
2/3
4
5
Layer 1 fibres
Tuft
Ca²⁺ AP
Apical dendrite
Oblique dendrites
Na⁺ AP
Basal dendrites
Thalamic afferents
50 mV
10 ms

Larkum, Zhu & Sakmann - Nature, 1999

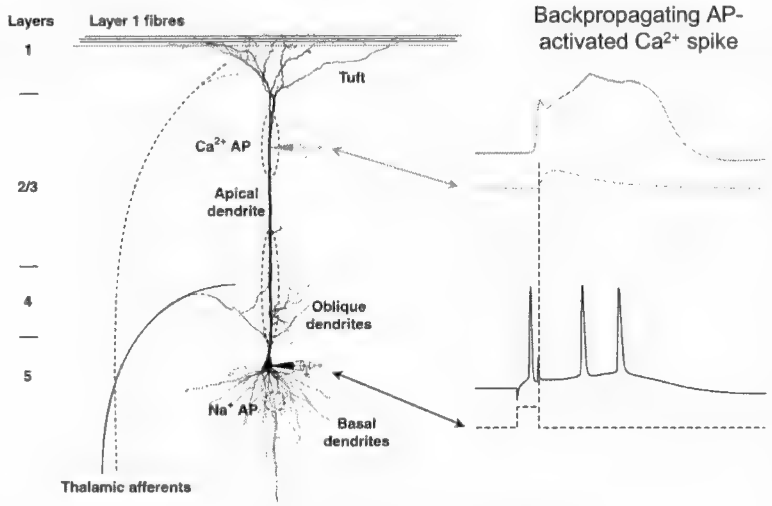
Interaction between backpropagating APs and distal Ca spikes


Back-propagating action potential only

Layers
1
2/3
4
5
Layer 1 fibres
Tuft
Ca²⁺ AP
Apical dendrite
Oblique dendrites
Na⁺ AP
Basal dendrites
Thalamic afferents
50 mV
10 ms

Larkum, Zhu & Sakmann - Nature, 1999

Interaction between backpropagating APs and distal Ca spikes


Layers
1
2/3
4
5
Layer 1 fibres
Tuft
Ca^{2+} AP
Apical dendrite
Oblique dendrites
Na^{+} AP
Basal dendrites
Thalamic afferents
Backpropagating AP-activated Ca^{2+} spike

Larkum, Zhu & Sakmann - Nature, 1999

Take-home messages

1. Synaptic integration is the way inputs are combined to generate output

2. The temporal and spatial pattern of inputs is critical to synaptic summation

3. Dendrites express voltage-gated channels which can promote spike backpropagation or trigger local spikes

4. Active dendrites generate functional subcompartments in the neuron

5. Synaptic integration in real neurons is more complex - and more powerful - than in simple, linear-summing units

Neuron *51*, 227–238, July 20, 2006 ©2006 Elsevier Inc. DOI 10.1016/j.neuron.2006.06.017

A Cooperative Switch Determines the Sign of Synaptic Plasticity in Distal Dendrites of Neocortical Pyramidal Neurons

Per Jesper Sjöström[1,2,*] and Michael Häusser[1,2]
[1] Wolfson Institute for Biomedical Research
[2] Department of Physiology
University College London
London WC1E 6BT
United Kingdom

Summary

Pyramidal neurons in the cerebral cortex span multiple cortical layers. How the excitable properties of pyramidal neuron dendrites allow these neurons to both integrate activity and store associations between different layers is not well understood, but is thought to rely in part on dendritic backpropagation of action potentials. Here we demonstrate that the sign of synaptic plasticity in neocortical pyramidal neurons is regulated by the spread of the backpropagating action potential to the synapse. This creates a progressive gradient between LTP and LTD as the distance of the synaptic contacts from the soma increases. At distal synapses, cooperative synaptic input or dendritic depolarization can switch plasticity between LTD and LTP by boosting backpropagation of action potentials. This activity-dependent switch provides a mechanism for associative learning across different neocortical layers that process distinct types of information.

Introduction

Learning and memory, as well as the refinement of cortical maps during development, are thought to be due to changes in the strength of individual synaptic connections between neurons (Katz and Shatz, 1996; Malenka and Nicoll, 1999). Hebb postulated that such synaptic strengthening occurs during persistent simultaneous activity in pre- and postsynaptic neurons (Hebb, 1949). At both hippocampal (Magee and Johnston, 1997) and neocortical (Markram et al., 1997b; Sjöström et al., 2001) synapses, repeated coincidence of EPSPs and postsynaptic action potentials (APs) results in long-term potentiation (LTP), thus satisfying the Hebbian postulate. Backpropagation of the AP into the dendrites appears to be a critical requirement for triggering changes in strength at many synapses (Linden, 1999). This coincidence detection depends at least in part on supralinear Ca^{2+} influx through NMDA receptors (Koester and Sakmann, 1998; Yuste and Denk, 1995) and on nonlinear summation mediated by voltage-gated dendritic channels (Magee and Johnston, 1997; Stuart and Häusser, 2001). However, the backpropagating AP (bAP) is typically strongly attenuated in distal dendrites of pyramidal neurons (Stuart et al., 1997; Stuart and Häusser, 2001). This suggests that bAP-dependent learning rules may be different at proximal and distal synapses (Froemke et al., 2005; Goldberg et al., 2002; Golding et al., 2002; Lisman and Spruston, 2005; Sourdet and Debanne, 1999). Neocortical layer 5 (L5) pyramidal neurons represent an excellent model system for addressing this issue, since these neurons have a large apical dendritic arbor that extends for hundreds of microns across several cortical layers, and since they receive two well-defined inputs, making synapses at different dendritic locations: the vast majority of L5-to-L5 synapses are on the proximal, basal dendritic tree (Markram et al., 1997a), while L2/3-to-L5 synapses are found on the apical as well as the basal dendrites (Thomson and Bannister, 1998). However, the consequences of synaptic location for plasticity exhibited by these two distinct inputs are not known.

Here we have investigated how plasticity rules for L2/3 and L5 inputs to L5 pyramidal neurons depend on the dendritic location of synaptic contacts and their interaction with the local excitable properties of the dendritic tree. We have used a combination of quadruple simultaneous patch-clamp recordings (Sjöström et al., 2001), dendritic patch-clamp recordings (Stuart and Sakmann, 1994), two-photon imaging (Koester and Sakmann, 1998; Yuste and Denk, 1995), and modeling of reconstructed neurons (Markram et al., 1997a; Redman and Walmsley, 1983) in order to test identical plasticity protocols at the two different unitary inputs and relate the outcome to the location of identified synaptic contacts and the dendritic Ca^{2+} signals associated with plasticity induction. We found that, surprisingly, the same AP-EPSP pairing protocol generates LTP at proximal L5-to-L5 connections, but LTD at distal L2/3-to-L5 connections, thus creating a gradient along the apical dendrite for the polarity of plasticity. We demonstrate that this gradient is defined by the efficacy of AP backpropagation to the respective synaptic sites, since dendritic depolarization or cooperative synaptic input can rescue LTP at distal synapses in a manner which is tightly linked to the boosting of AP backpropagation into the distal dendritic tree. We also show that these electrical interactions are reflected in terms of dendritic Ca^{2+} signals. Our findings provide a dendritic mechanism by which inputs in deeper cortical layers can regulate plasticity at inputs in more superficial layers.

Results

We compared plasticity of synaptic inputs mediated by two distinct input pathways in L5 pyramidal neurons by making simultaneous recordings from L5 neurons and presynaptic L5 and L2/3 pyramidal neurons. We used the same plasticity protocol at both synaptic connections, consisting of pairing five EPSPs and APs at 50 Hz. This frequency is below the critical frequency for evoking Ca^{2+} spikes in pyramidal cell dendrites (Larkum et al., 1999a). Pairing EPSPs and APs at L5-to-L5 connections produced robust LTP (Figure 1A; after/before 140% ± 6%, n = 34; Markram et al., 1997b; Sjöström et al., 2001). In contrast, when the identical pairing protocol was applied at unitary L2/3-to-L5 connections, this

---

*Correspondence: j.sjostrom@ucl.ac.uk

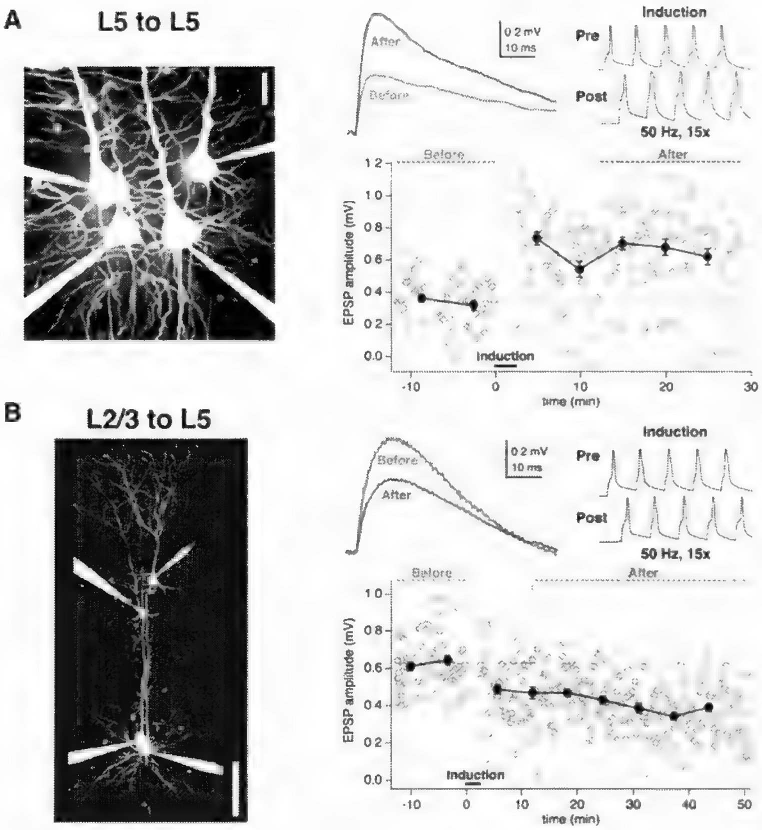


Figure 1. AP-EPSP Pairing Evokes LTP at an L5-to-L5 Connection, but LTD at an L2/3-to-L5 Connection

(A) AP-EPSP pairing at a unitary L5-to-L5 connection resulted in LTP. After the baseline period (blue dotted line and blue trace, top middle inset), pre- and postsynaptic 50 Hz AP trains were paired (pre leading before post by 10 ms; top right inset). The post-pairing responses (red dotted line and red trace, top middle inset) were significantly potentiated (after/before = 195%, $p < 0.001$; EPSP rise time = 1.9 ms), in agreement with prior studies (Markram et al., 1997b; Sjöström et al., 2001). (Left) Illustration of recording configuration for L5-to-L5 pairs.

(B) The same conditions and induction protocol as in (A) in this L2/3-to-L5 paired recording, however, resulted in significant LTD (after/before = 63%, $p < 0.001$; EPSP rise time 3.5 ms). (Left) Recording configuration.

did not produce LTP (106% ± 9%, $n = 19$; $p < 0.001$ compared to L5-to-L5 connections). Indeed, for many L2/3-to-L5 pairs this induction protocol resulted in LTD (Figure 1B).

Identifying the Dendritic Location of Synapses

Synaptic contacts of L5-to-L5 and L2/3-to-L5 connections tend to be made at different locations in the dendritic tree of L5 pyramidal neurons (Markram et al., 1997a; Thomson and Bannister, 1998). We used a combination of computer modeling and direct visualization of contacts to identify the location of synaptic contacts in our connected pairs. Cable theory predicts that EPSP rise time should be correlated with dendritic location of the active synaptic contacts (Rall, 1967). We made detailed compartmental models based on three-dimensional reconstructions of biocytin-filled L5 pyramidal neurons from our paired recordings (Figure 2A; see the Experimental Procedures). Results from simulations using such a model are shown in Figures 2B and 2C, which allow us to quantitatively predict the relationship between synaptic location and somatic EPSP rise time under the conditions of our experiments for the entire range of possible synaptic locations.

We validated this approach experimentally by directly identifying putative synaptic contacts during some recordings by imaging contacts made by the presynaptic axon with the postsynaptic target using two-photon laser-scanning microscopy (2PLSM) of neurons filled with fluorescent dye (see Figure S1 in the Supplemental Data). Alternatively, we filled neurons with biocytin and identified putative synaptic contacts using bright-field microscopy of fixed and histologically processed neurons (Figure 2A). Across the population, the distance from the soma of visually identified putative synapses was tightly correlated with somatic EPSP rise time (Figure 2E). This provides a model-independent index relating somatic EPSP rise time and synaptic location, which closely agrees with the prediction of our simulations (Figure 2C). Finally, we used this index to verify that—as predicted on the basis of anatomy (Markram et al., 1997a; Thomson and Bannister, 1998)—the location of functional L2/3-to-L5 synapses (rise time 3.0 ± 0.4 ms; $n = 19$) was on average significantly farther from the soma than that of L5-to-L5 synapses (2.2 ± 0.2 ms, $n = 34$, $p < 0.05$).

The Sign of Plasticity Depends on Synaptic Location

Next, we used EPSP rise time as an index of synaptic location to determine if location is correlated with the outcome of the plasticity protocol. Across all unitary L5-to-L5 and L2/3-to-L5 connections, we found a clear relationship between the sign of plasticity and the synaptic location (estimated by EPSP rise time). Thus, proximal inputs with fast-rising EPSPs exhibited classical Hebbian LTP, whereas distal inputs with slow-rising EPSPs underwent LTD (Figure 3). In a subset of experiments, where we also identified putative synaptic contacts (Figure S1), we could demonstrate directly that the sign of plasticity depended on the distance of synaptic contacts from the soma (inset, Figure 3). We confirmed and extended these results by activating inputs at a range of distances from the soma using extracellular

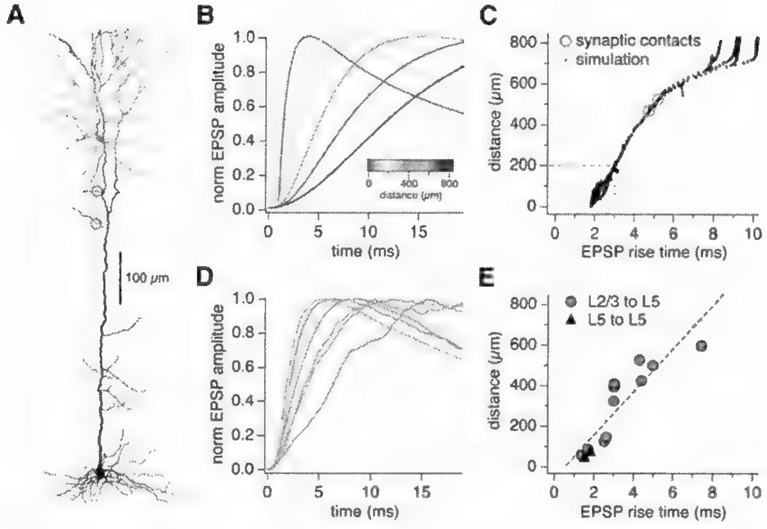


Figure 2. Identifying Synaptic Locations from EPSP Rise Time and Morphology

(A) Reconstruction of an L2/3-to-L5 connected pair (blue circles, putative synaptic contacts).

(B) Using a computer model of the reconstructed L5 neuron in (A), calibrated to the actual EPSP rise time (5.0 ms) of this paired recording by assuming equal contribution of the two putative synaptic contacts, we simulated the activation of synapses in different locations of the dendritic tree. This gave rise to EPSPs with different rise times (color-coded according to synapse location).

(C) Based on this simulation, we predicted the relationship between distance and somatic EPSP rise time for many synaptic locations. Dashed line: 3 ms cut-off that defines proximal versus distal inputs (see text).

(D) In several paired recordings, putative synaptic contacts were identified (cf. Figure S1), which enabled the verification of the simulation (cf. [B]).

(E) The mean location of putative synaptic contacts correlated well (r = 0.89) with the somatic EPSP rise time in L5-to-L5 (black triangles) and L2/3-to-L5 (closed blue circles) paired recordings (n = 13), thus corroborating the model prediction (cf. [C]).

stimulation (Figure 3). While overall the degree of plasticity was statistically different between L2/3-to-L5 and L5-to-L5 pathways, the overlap in the distribution of synaptic locations and the strong correlation with rise time (Figure 3) indicate that synaptic location in the dendritic tree, rather than input origin per se, is the major determinant of plasticity.

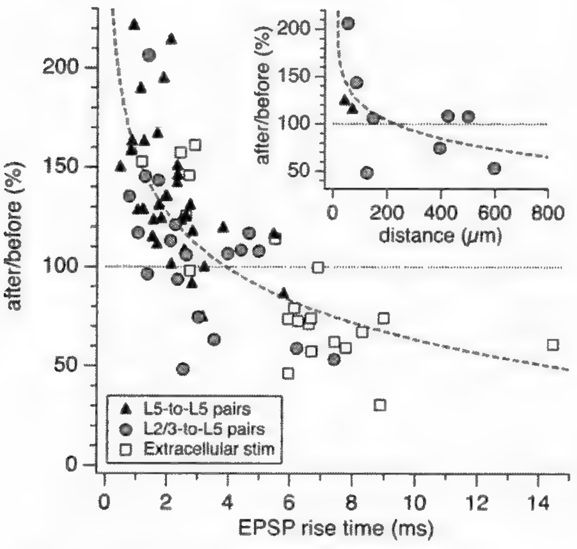


Figure 3. Synaptic Location Determines the Sign of Plasticity

AP-EPSP pairing in L5-to-L5 pairs (closed black triangles, n = 34), L2/3-to-L5 pairs (closed blue circles, n = 19), and extracellular stimulation experiments (open black squares, n = 21) resulted in LTP of proximal inputs (somatic EPSP rise time < 3 ms; 137% ± 4.9%, n = 46), but evoked LTD of distal input (somatic EPSP rise time > 3 ms; 80.2% ± 4.5%, n = 28, $p < 0.001$). (Inset) In agreement with this, the sign of plasticity was inversely correlated with average distance of putative synaptic contacts (n = 11; also see Figure S1). Dashed red lines are power law fits.

Converting LTD to LTP at Distal Synapses

To determine if distal inputs, which tend to exhibit LTD, can be converted to exhibit LTP, we selectively activated distal synaptic inputs near the apical tuft of L5 neurons using extracellular stimulation, restricting ourselves to the most distal inputs with a rise time >3 ms (corresponding to >200 μm from the soma, cf. Figure 2C). Since it takes an AP several milliseconds to backpropagate the entire length of the apical dendrite (Stuart and Sakmann, 1994), it is conceivable that this delay might push pre- and postsynaptic spike pairings out of the temporal window for LTP and into that of LTD (Sjöström et al., 2001). To test this hypothesis, we modified our pairing protocol to evoke postsynaptic spikes 5 ms (instead of 10 ms) after extracellular stimulation near the apical tuft. This compensation of the temporal delay did not switch distal LTD to LTP (data not shown; after/before = 89% ± 5%, EPSP rise time 5.2 ± 0.6 ms, amplitude 0.4 ± 0.1 mV, n = 4). This result indicates that the delay due to AP backpropagation does not underlie the location dependence of plasticity shown in Figure 3, consistent with the observation that precise timing of pre- and postsynaptic spikes has relatively little impact on plasticity evoked by correlated 50 Hz firing (Sjöström et al., 2001).

However, for distal inputs, recruiting more inputs by increasing stimulation strength resulted in large EPSPs that potentiated when paired with APs at 50 Hz (Figures 4A and 4C). To directly verify that this cooperativity rule holds for identified unitary distal inputs, we made paired recordings between L2/3 and L5 pyramidal neurons exhibiting distal synaptic connections, which normally do not potentiate when paired at 50 Hz with APs (Figure 3). When such connections were subjected to the conventional pairing protocol in combination with strong extracellular stimulation to boost the unitary EPSP, these connections exhibited LTP (Figures 4B and 4C). Pooling data from all distal inputs (both paired recordings as well

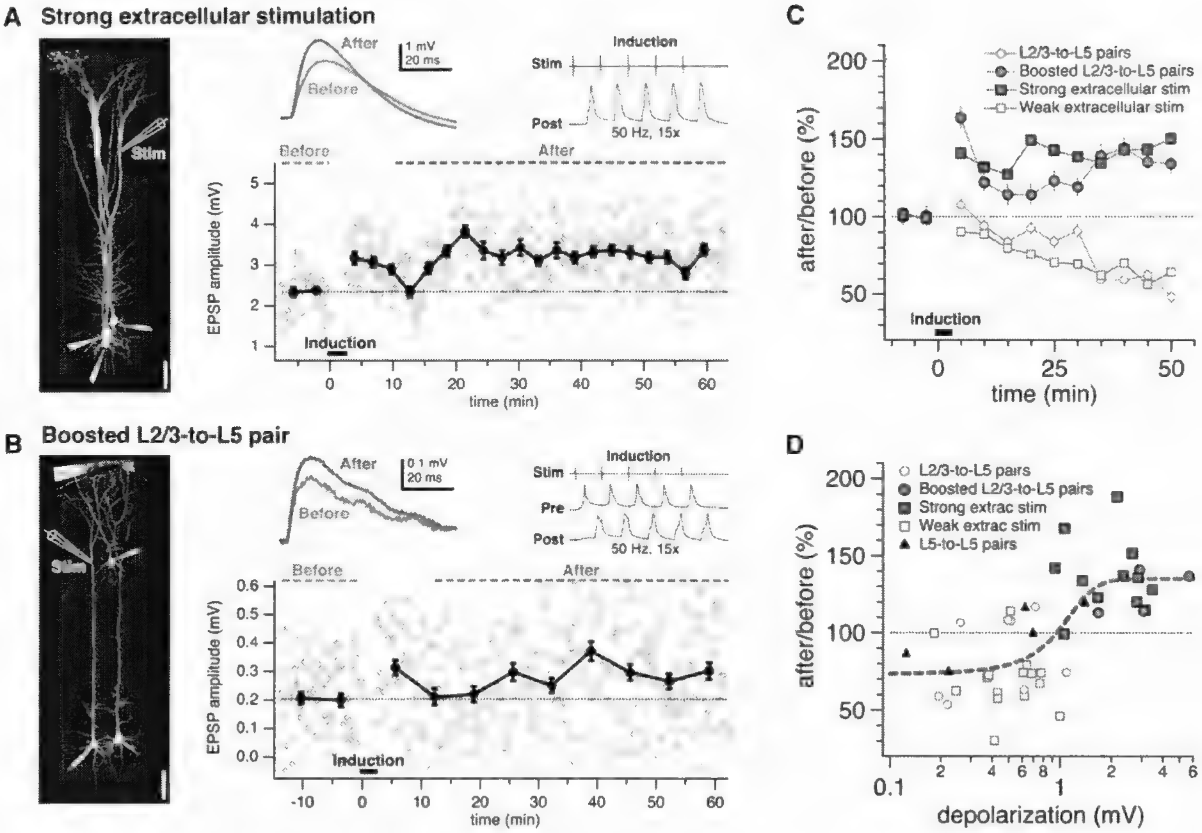


Figure 4. LTP Induction at Distal Synapses Requires Cooperativity among Inputs

(A) Strong extracellular stimulation near the distal dendritic tuft (left) recruited a slow, large EPSP (rise time 5.2 ms) that significantly potentiated when paired with APs at 50 Hz (137%, $p < 0.001$). (Left) Configuration permitted multiple parallel recordings.

(B) A distal L2/3-to-L5 unitary input (rise time 4.2 ms) that was boosted by extracellular stimulation during the induction (compound EPSP 5.7 mV) potentiated significantly after AP-EPSP pairing (137%, $p < 0.001$). (Left) Stimulation electrode placement is illustrated schematically.

(C) Using extracellular stimulation, pairing small distal EPSPs with APs resulted in LTD (closed red squares, 70% ± 5%, $n = 15$). Recruiting more inputs by increasing extracellular stimulation strength, however, resulted in large EPSPs that potentiated robustly when paired with APs (open red squares, 137% ± 7%, $n = 12$; $p < 0.001$), demonstrating a cooperativity requirement for distal LTP. Similarly, although distal unitary connections typically depressed (open blue circles, 86% ± 9%, $n = 8$; cf. Figure 3), boosted distal unitary connections (cf. [B]) significantly potentiated (126% ± 7%, $n = 4$, $p < 0.05$). (D) For distal inputs, LTD (80% ± 5%, $n = 25$) was observed below a threshold somatic EPSP amplitude (<1.0 mV). Above this threshold, distal plasticity switched to LTP (126% ± 7%, $n = 19$; $p < 0.001$). Data is pooled from L5-to-L5 pairs (triangles), L2/3-to-L5 pairs (open blue circles), boosted L2/3-to-L5 pairs (closed blue circles) and extracellular stimulation experiments (open and closed squares). Dashed red line is best fit sigmoid. Only experiments with EPSP rise time > 3 ms (cf. Figure 2) were included in (C) and (D).

as extracellular stimulation), we found that AP-EPSP pairing of distal inputs above a threshold amplitude reliably resulted in LTP (Figure 4D; sigmoid $x_{half} = 1.0$ mV; sliding t test minimum 1.06 mV; see Experimental Procedures). In other words, the sign of distal plasticity was switched from negative (LTD) to positive (LTP) once a sufficient number of distal inputs cooperated (Figure 4D).

The cooperativity requirement for distal LTP might be due to a need for local dendritic spikes, as was recently found in CA1 pyramidal neurons (Golding et al., 2002), in which case pairing with somatic APs may not be required for LTP of large distal EPSPs. Activating distal EPSPs at 50 Hz in the absence of postsynaptic somatic spiking, however, did not produce LTP, but surprisingly resulted in LTD (Figures 5A and 5B). Postsynaptic 50 Hz spiking in the absence of presynaptic stimulation, on the other hand, did not change synaptic strength (Figure 5B). These results show that, in the absence of postsynaptic spiking, activation of large distal inputs at 50 Hz (but not at 0.1 Hz, our baseline firing rate) results in LTD. We hypothesized that this form of LTD—which apparently depends on pairing presynaptic activity with subthreshold postsynaptic depolarization—was similar to the endocannabinoid-dependent, depolarization-induced LTD previously described at proximal L5-to-L5 synapses (Sjöström et al., 2004). In keeping with this view, the endocannabinoid CB1 receptor blocker AM251 abolished this form of distal LTD (Figure 5B) (Sjöström et al., 2004). Taken together, these results show that in L5 neurons, distally evoked EPSPs undergo LTP when a dual requirement for coincidence with APs and cooperativity among inputs is satisfied. For inputs that are activated at 50 Hz, failure to satisfy either requirement results in LTD.

Boosting of bAPs Underlies the Switch between Distal LTD and LTP

Can the cooperativity requirement for distal LTP be explained by the boosting of failing bAPs? Backpropagating APs may not be sufficient to unblock NMDA

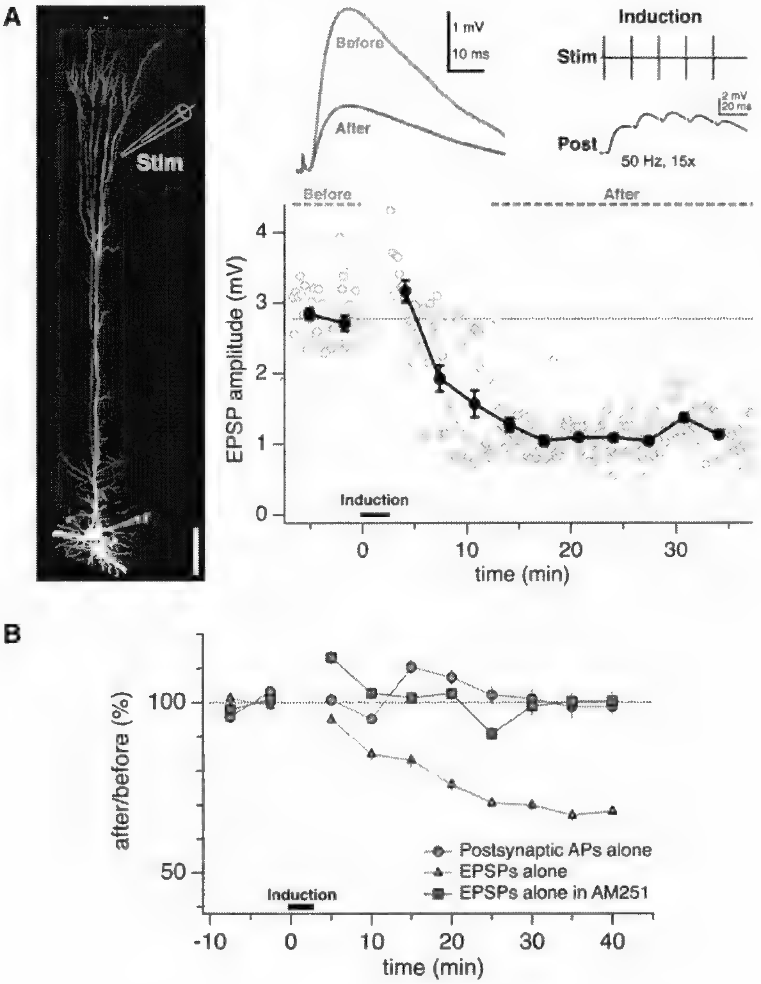


Figure 5. Distal LTD is Endocannabinoid-Dependent

(A) Extracellular 50 Hz stimulation of distal inputs (rise time 3.7 ms) in the absence of somatic spiking resulted in depression (41%, $p < 0.001$). (Left) Recording configuration.

(B) Unpaired distal EPSPs evoked LTD (green triangles, 71% ± 6%, n = 10; cf. Golding et al., 2002) that was abolished by the endocannabinoid CB1 receptor blocker AM251 (Gatley et al., 1996) (red squares, 99% ± 3%, n = 6, $p < 0.01$). Unpaired spike trains (blue circles, 102% ± 6%, n = 10) did not evoke plasticity.

receptors and induce LTP for small distal inputs (Sjöström and Nelson, 2002), since APs backpropagate decrementally (Larkum et al., 1999a, 2001; Stuart et al., 1997; Stuart and Häusser, 2001; Williams and Stuart, 2000), but bAPs may still depolarize the distal dendrites sufficiently to help evoke LTD (Artola et al., 1990; Sjöström et al., 2004). Sufficiently large distal EPSPs, on the other hand, may potentiate because they provide enough dendritic depolarization to boost failing bAPs (Larkum et al., 2001; Stuart and Häusser, 2001; Williams and Stuart, 2000). If so, then small distal EPSPs should potentiate if the bAP is boosted by direct dendritic depolarization (Larkum et al., 2001; Stuart and Häusser, 2001; Williams and Stuart, 2000). In agreement with this, we found that brief subthreshold dendritic current injection in conjunction with the pairing of APs and small distal EPSPs produced LTP (Figure 6). Somatic subthreshold depolarization of the same amplitude and duration, however, did not evoke LTP of small distal EPSPs (Figure 6B), showing that dendritic, but not somatic, depolarization determines the sign of distal plasticity.

To determine whether the efficacy of AP backpropagation is linked to distal LTP induction, we examined how LTP-promoting conditions influenced AP backpropagation. Dendritic depolarization boosted the amplitude and increased the half-width of dendritic bAPs (Figure 7A) in a manner that increased with distance from the soma (Figure 7B), consistent with previous findings (Larkum et al., 2001; Stuart and Häusser, 2001; Williams and Stuart, 2000). To examine bAP boosting in multiple locations on the same neuron and in thin distal dendrites where dendritic recordings are not possible, we used 2PLSM imaging of bAP-induced Ca^{2+} responses (see Experimental Procedures). Supralinear Ca^{2+} signals were generated throughout the distal apical dendritic tree by pairing APs with dendritic depolarizing current injection (supralinearity = 268% ± 68%, n = 19, $p < 0.001$; Figure 7C) or with large EPSPs (100% ± 30%, n = 37, $p < 0.001$; Figure 7B), suggesting that boosting of the otherwise-decaying bAP (Larkum et al., 2001; Stuart and Häusser, 2001; Williams and Stuart, 2000) (see Figure S2) occurs over large parts of the distal dendritic tree. The Ca^{2+} signal, due to large EPSPs in the absence of APs, was minimal (Figure 7D), consistent with the need for somatic APs in LTP induction (Figure 5). As with dendritic recordings (Figure 7B), the boosting of the Ca^{2+} signal was more effective in the distal regions of the dendritic tree (Figure 7E), where bAP failure is most prominent (Larkum et al., 1999a, 2001; Stuart and Häusser, 2001; Williams and Stuart, 2000) (cf. Figure S2). Pairing of unboosted bAPs and small distal

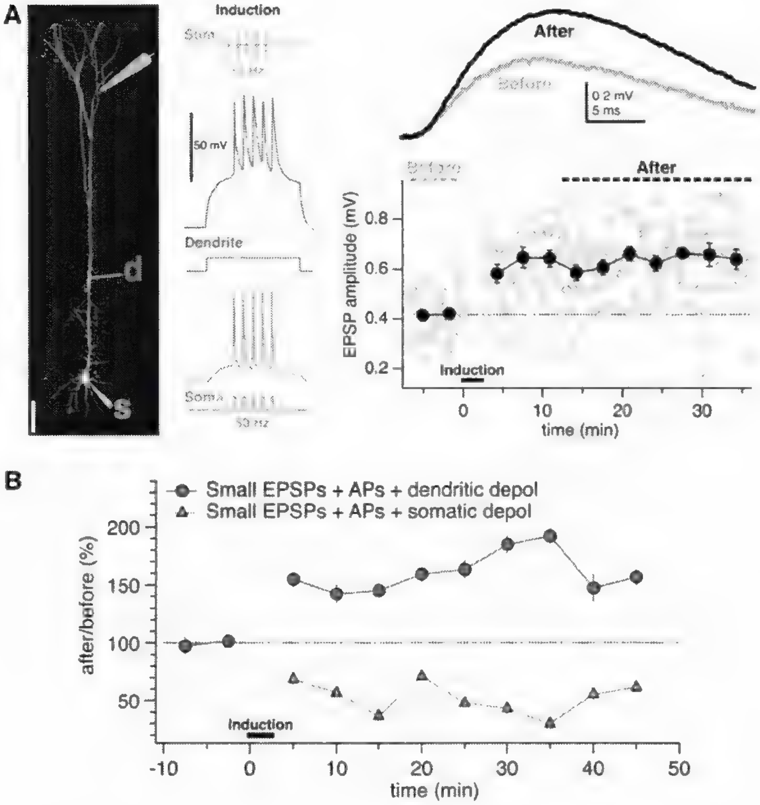


Figure 6. Dendritic Depolarization Switches Distal Plasticity from LTD to LTP

(A) AP-EPSP pairing in conjunction with a brief subthreshold dendritic current injection (0.4 nA, 200 ms, red trace) resulted in LTP of a small, distal input (rise time 4.9 ms). (Left) Recording configuration, also showing sample LTP induction traces (stimulus electrode, green; dendrite, red, "d;" soma, blue, "s").

(B) Pairing of APs and EPSPs in the presence of a brief subthreshold dendritic current injection (247 ± 22 μm from the soma) evoked LTP of small distal inputs (<1 mV; red circles; 163% ± 7%, n = 5). The corresponding amount of subthreshold current injected into the soma, however, resulted in LTD (blue triangles; 88% ± 8%, n = 3). The subthreshold dendritic (0.35 ± 0.03 nA) and somatic (0.31 ± 0.01 nA) current injections were not significantly different (p = 0.36). These results demonstrate that dendritic, but not somatic, depolarization switches distal plasticity from LTD to LTP (p < 0.001).

EPSPs (<1.01 mV; Figure 4C), however, did not result in supralinear Ca^{2+} influx (−2% ± 8%, n = 9; EPSP amplitude 0.37 ± 0.2 mV; data not shown), consistent with the requirement for sufficiently large EPSPs in bAP boosting and in LTP (Figures 4C and 4D). In five cells, we increased the extracellular stimulation strength stepwise to determine the threshold for AP-EPSP-induced supralinear Ca^{2+} signals. This threshold (sigmoid x_{half} = 1.1 ± 0.32 mV) matched that of LTP induction (see above and Figure 4D). Taken together, these results identify a depolarization threshold for distal LTP (Figures 4C and 4D) that is due to a requirement for boosting of the bAP.

LTP at Proximal Synapses also Requires Reliable AP Backpropagation

If the cooperativity requirement of distal LTP results from the need to boost failing bAPs, then weakening AP propagation into basal dendrites should in turn affect plasticity of proximal inputs in a converse manner. We therefore examined the effect of delivering strong, brief postsynaptic hyperpolarization on plasticity at proximal connections during our 50 Hz pairing protocol. First, we investigated whether this manipulation indeed impairs backpropagation into basal dendrites during the pairing protocol, using dendritic Ca^{2+} signals as an index of backpropagation. Strong, brief hyperpolarization caused a reduction in the dendritic Ca^{2+} signal associated with backpropagating APs, and this reduction increased progressively along both the basal and apical dendrites with distance from the soma, while leaving the axonal signal unaffected (Figures 8A and 8B). These results are thus consistent with impaired dendritic backpropagation due to somatic hyperpolarization. Next, we showed that AP-EPSP pairing during the hyperpolarization abolished LTP at proximal connections (Figures 8C and 8D; cf. Markram et al., 1997b; Sjöström et al., 2001). This result is in keeping with the view that reliability of AP backpropagation determines plasticity in L5 neurons, regardless of synapse location.

While somatic hyperpolarization abolishes LTP due to AP-EPSP pairing at proximal synapses (Figures 8A, 8C, and 8D) (Sjöström et al., 2001), it does not switch plasticity from LTP to LTD. This is perhaps because even failing bAPs still provide sufficient depolarization at proximal synapses to produce a balance between LTP and LTD, which results in no net plasticity. If so, then unpaired proximal inputs should show LTD, just like distal inputs (Figure 5). Consistent with this view, strong 50 Hz extracellular stimulation of proximal inputs resulted in LTD (Figure S3). This LTD required strong stimulation, as it is not observed with unitary connections (Markram et al., 1997b; Sjöström et al., 2003). The dual need for presynaptic high-frequency activity and strong postsynaptic depolarization to evoke LTD at proximal synapses (Figure S3) is similar to the distal LTD observed in the absence of bAPs (Figure 5), and it is also in keeping with prior results showing that LTD requires presynaptic activity paired with sufficiently strong postsynaptic depolarization (Sjöström et al., 2003, 2004). Taken together, these results confirm the key importance of backpropagation in regulating synaptic plasticity, and they suggest that failing or absent bAPs affect plasticity of proximal and distal synapses in a similar manner.

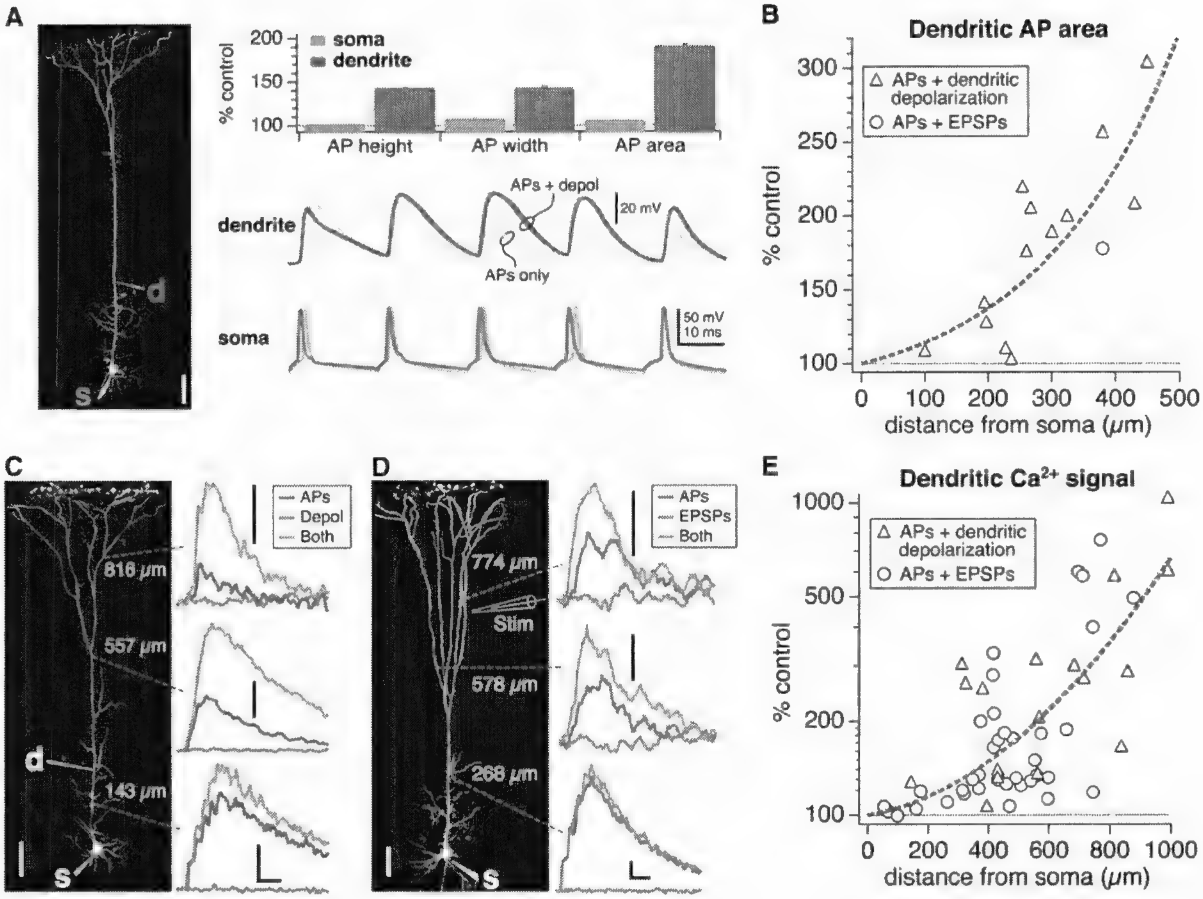


Figure 7. Conditions that Evoke Distal LTP also Boost bAPs

(A) Amplitude, width, and area of bAPs were increased by a 0.4 nA subthreshold dendritic current injection (red bars, top, and red traces in middle). At the soma, the effect on the AP was moderate (blue bars, top, and blue traces at bottom). Traces with 0.4 nA subthreshold current injection are represented with darker colors than control traces (middle, "APs + depol," and bottom). (Left) Dendritic recording ("d") 301 μm from the soma ("s").

(B) The amount of depolarization-induced bAP boosting increased with dendritic recording distance from the soma (Larkum et al., 1999a, 2001; Stuart and Häusser, 2001). The boosting occurred irrespective of the source of the depolarization (dendritic current injection, triangles; synaptic activation, circles; n = 13 cells total).

(C) (Left) L5 pyramidal cell filled with Fluo-5F and Alexa Fluor 594. (Right) Line scans at three dendritic locations reveal supralinear Ca^{2+} signals (green) when pairing APs with a subthreshold 200 ms long dendritic current injection (0.4 nA, 260 μm from soma, "d") in distal dendrites, compared to the arithmetic sum of the APs (red) and the depolarization alone (blue). Note that boosting is widespread in distal dendrites. In this cell, the simultaneous dendritic recording measured a bAP area boosting of 175% ± 0.4%; compare (B). Scale bars, 100 μm, 20% dG/R, 50 ms.

(D) APs paired with distal large EPSPs also resulted in supralinear Ca^{2+} signals (green) in distal dendrites, as compared with APs (red) and EPSPs (blue). As with plasticity experiments (Figure 4), EPSPs (3.9 mV; rise time 3.2 ms) were evoked by relatively strong extracellular stimulation near the dendritic tuft (stimulation electrode position 740 μm). Note that boosting is not limited to locations close to the stimulation electrode. Scale bars, 2% dG/R, and see (C).

(E) The Ca^{2+} signal boosting due to pairing of APs with depolarization (triangles; 18 positions, four cells) or with large EPSPs (circles; 37 positions, ten cells) increased with distance from the soma, indicating that sufficient dendritic depolarization boosts bAPs throughout the distal apical dendritic tree (Larkum et al., 1999a, 2001; Stuart and Häusser, 2001).

Discussion

Our results reveal two principles for the organization of synaptic plasticity rules in neocortical L5 pyramidal neurons. First, there is a graded change in the sign and magnitude of plasticity with distance from the soma: whereas high-frequency correlated firing evokes Hebbian LTP at proximal synapses, it surprisingly results in LTD of unitary distal inputs. Second, a voltage-dependent dendritic switch governs plasticity at distal synapses, such that dendritic depolarization switches between LTD and LTP given identical patterns of pre- and postsynaptic activity. These principles are both governed by the amplitude of the backpropagating AP at the sites of the active synapses. Our study thus provides some of the most direct evidence to date that the efficacy of AP backpropagation determines the magnitude and sign of synaptic plasticity in dendrites. The resulting model of plasticity in L5 pyramidal neurons is summarized in schematic form in Figure 9.

Inverted Sign of Plasticity at Distal Synapses

We show that the same standard AP-EPSP pairing protocol can produce LTP at proximal synapses and LTD at distal synapses. To our knowledge, such an inversion of the sign of plasticity depending on dendritic distance

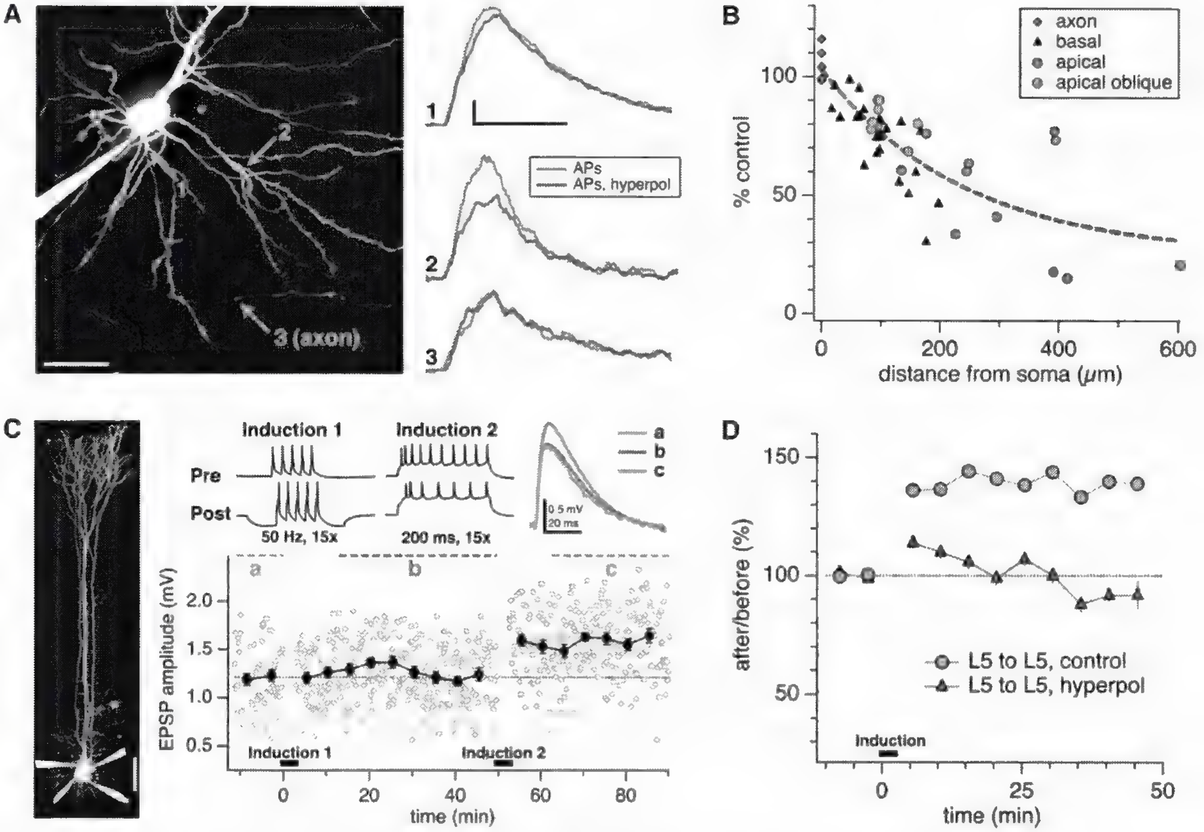


Figure 8. Somatic Hyperpolarization Reduces Basal Dendritic AP-Evoked Ca^{2+} Signals and Abolishes Proximal LTP

(A) (Left) L5 pyramidal cell filled with Fluo-5F and Alexa Fluor 594. (Right) Line scans in basal dendritic locations 1 (21 μm from soma) and 2 (92 μm) reveal reduced Ca^{2+} signals for APs paired with hyperpolarization (red; −0.34 nA, 200 ms, −23 mV max deflection) compared with APs alone (blue). Note the difference in the effect at locations 1 and 2, as well as the absence of an effect in the axon (location 3, 147 μm). Scale bars, 50 μm, 20% dG/R, 200 ms.

(B) AP-evoked Ca^{2+} signals in dendrites, but not in axons, are reduced more by hyperpolarization as the distance from the soma increases. This data is consistent with a hyperpolarization-induced reduction in the reliability of AP propagation into the basal ($n = 25$) as well as the apical ($n = 12$) and apical oblique ($n = 7$) dendritic arbors, but not into the axon ($n = 6$). Data was collected from five cells. Axonal line scans, indicated at origin for clarity, were acquired 129 ± 14 μm from soma.

(C) In a connected L5-to-L5 pair, strong postsynaptic hyperpolarization during 50 Hz pairing (Induction 1; −0.41 nA, 200 ms) abolished LTP (Sjöström et al., 2001). The same connection subsequently potentiated due to a strong LTP induction protocol (Induction 2; 0.7 nA, 200 ms; 130%, $p < 0.001$), demonstrating that its ability to express LTP was not compromised. Scale bar, 100 μm.

(D) Somatic hyperpolarization blocked LTP in L5-to-L5 pairs (red triangles; $n = 6$) as compared with controls without hyperpolarization (green circles; $n = 34$).

has not been demonstrated previously. A recent study in L2/3 pyramidal neurons (Froemke et al., 2005) showed that the temporal window of spike timing-dependent plasticity is different at proximal and distal synapses, leading to different degrees of potentiation or depression, but did not demonstrate a difference in sign of plasticity using the same protocol. Our results indicate that in L5 pyramidal neurons, synaptic location is a critical determinant not simply of the extent, but also of the polarity of plasticity at each given synapse. This finding may help to explain some of the notorious variability in the outcome of plasticity experiments, where synaptic location is normally not rigorously controlled.

A Dendritic Switch for Regulating Plasticity at Distal Synapses

We demonstrate that distal synapses—which normally undergo LTD when paired with bAPs—undergo LTP in the presence of either cooperative synaptic input or dendritic depolarization. We show that this dendritic switch for plasticity is directly linked to the boosting of bAPs by dendritic depolarization. Magee and Johnston (1997) demonstrated that reliable AP backpropagation is critical for LTP induction in CA1 pyramidal neurons when pairing APs and EPSPs. Several studies have also shown that such AP-EPSP pairing can boost AP backpropagation in a manner that depends on nonlinear recruitment of dendritic voltage-gated Na^+ channels (Bernard and Johnston, 2003; Magee and Johnston, 1997; Stuart and Häusser, 2001; Watanabe et al., 2002), but how such boosting is converted into changes in synaptic strength is not well understood, particularly at distal synapses. We show that the same stimuli that boost bAPs also promote the conversion of LTD to LTP at distal synapses, with a close correspondence in the threshold amplitude of EPSPs required to boost backpropagation and trigger the plasticity switch. This finding directly links the mechanism for electrical

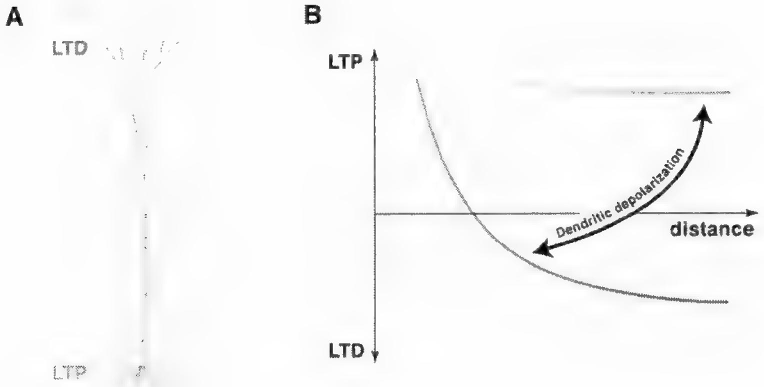


Figure 9. A Dendritic Switch Providing a Learning Rule for Distal Dendritic Long-Term Plasticity

(A) Proximal inputs onto an L5 pyramidal cell are relatively more prone to LTP (green) than distal dendritic inputs, which tend to undergo LTD (red).

(B) Sufficient depolarization of the apical dendrite switches distal plasticity from LTD to LTP. Dendritic depolarization even relatively close to the soma can switch the mode of plasticity at the distal apical dendrite.

coincidence detection in dendrites to synaptic plasticity. We further demonstrate that activation of the plasticity switch is linked to highly supralinear increases in dendritic Ca^{2+}, which—given the well-established role of postsynaptic Ca^{2+} elevations in the induction of long-term plasticity (Cummings et al., 1996; Hansel et al., 1997)—may itself influence synaptic plasticity. Our results also provide a mechanism for cooperativity in strengthening distal synaptic inputs, a well-known feature of LTP (Debanne et al., 1996; Hebb, 1949; Levy and Steward, 1979; McNaughton et al., 1978).

Importantly, this plasticity switch is a specifically dendritic feature, since it relies on modulation of AP backpropagation (Hoffman and Johnston, 1999; Stuart et al., 1997). In principle, the mechanisms we describe may operate in any neuron exhibiting decremental AP backpropagation in conjunction with the expression of dendritic voltage-gated Na^+ channels (Bernard and Johnston, 2003; Stuart and Häusser, 2001). Neurons with different dendritic branching patterns or distributions of dendritic voltage-gated channels, such as hippocampal CA1 (Golding et al., 2002; Magee and Johnston, 1997) or neocortical L2/3 pyramidal neurons (Froemke et al., 2005), may exhibit a different balance and spatial distribution of learning rules. In CA1 pyramidal neurons, for example, somatic APs have little or possibly no impact on LTP of distal inputs, while local dendritic spikes play a prominent role (Golding et al., 2002). In L5 pyramidal neurons, however, somatically evoked bAPs play a critical role in the induction of LTP (Figure 5), and reliability of AP backpropagation determines the sign of plasticity of distal inputs (Figures 6 and 7). This difference may be related in part to the much higher density of A-type potassium channels in distal CA1 dendrites (Hoffman et al., 1997) as compared with L5 dendrites (Bekkers, 2000; Korngreen and Sakmann, 2000).

In the intact brain, where background synaptic activity is more prominent (Destexhe et al., 2003), the transition point between LTP and LTD along the apical dendrite may be tuned by the balance between excitation and inhibition and by their relative distribution along the dendrite. The additional shunt expected from background activity should influence the electrotonic structure of the dendritic tree as well as AP backpropagation (Destexhe et al., 2003) and thus may also affect the activity threshold for the transition between LTP and LTD. Nevertheless, boosting of AP backpropagation by synaptic input has recently been demonstrated in apical dendrites of L2/3 neocortical pyramidal neurons in vivo (Waters and Helmchen, 2004). This suggests that both the efficacy of backpropagation and the level of synaptic input in vivo are within the range that allows effective gating of the cooperative switch, and therefore that this mechanism can be used to regulate the sign of distal dendritic plasticity in vivo.

Functional Implications

Plasticity is notoriously sensitive to the details of the experimental protocol. We have studied the effect of one particular pairing protocol on synaptic efficacy in young animals. Using other induction frequencies or animal ages may lead to other outcomes; for example, different dendritic mechanisms may be engaged. With these provisos, our results have several important implications for the regulation of plasticity and the distribution of synaptic weights in the dendritic tree. Recent theoretical work has suggested that inverse plasticity rules may help equalize synaptic efficacy across the dendritic tree (Goldberg et al., 2002; Rumsey and Abbott, 2004). However, the propensity for proximal LTP and distal LTD that we have found (Figure 9) may instead lead to a distribution of strong proximal and weak distal inputs. In agreement with this, such a distribution of synaptic efficacies was recently described in L5 neurons (Williams and Stuart, 2002). These relatively weak distal inputs thus must act in tight synchrony to trigger somatic AP output, potentially aided by the generation of dendritic spikes (Golding and Spruston, 1998; Larkum et al., 1999b; Schiller et al., 1997; Williams and Stuart, 2002, 2003).

Functionally, the predisposition of distal synapses for LTD may help ensure the requirement for tight coincident activation of distal inputs in evoking somatic spike output by keeping distal inputs small relative to the threshold for spike generation (Schiller et al., 1997; Williams and Stuart, 2002). In addition, the existence of distal LTD even in the presence of somatic spiking can weaken unreliable and poorly correlated distal inputs by penalizing those distal connections that intermittently fail to satisfy the cooperativity criterion for LTP (discussed below). Even quite strong somatic hyperpolarization cannot promote a similar form of LTD at proximal synapses (Figure 8), highlighting the location-specific differences in plasticity between proximal and distal inputs.

Cooperative LTP, on the other hand, may reward inputs that act in synchrony and can thus promote associative learning (Hebb, 1949; Levy and Steward, 1979; McNaughton et al., 1978). Since we have shown that distal LTP can be gated by relatively proximal depolarization, our results demonstrate a mechanism for associative learning not just within but also across neocortical layers. This possibility is intriguing since input functionality varies with location in L5 neurons: proximal inputs are mostly of local origin (Markram et al., 1997a; Thomson and Bannister, 1998), whereas L1 inputs onto the apical tuft provide feedback from higher cortical regions (Coogan and Burkhalter, 1990) and thus carry distinct information compared to local inputs. In principle, proximal L2/3-to-L5 connections could therefore gate LTP and "instruct" learning at more distal L1-to-L5 inputs. The state-dependent switch between LTD and LTP at distal synapses may thus both regulate synaptic efficacy as a function of network activity and provide a means for associating and storing information carried by synchronous activity in different cortical layers.

Experimental Procedures

Slice Preparation and Electrophysiology

Neocortical brain slices were prepared according to standard techniques (Sjöström et al., 2001). All procedures conformed to the UK Animals (Scientific Procedures) Act 1986. Two- to three-week-old rats were anesthetized with isoflurane and decapitated, and the brain was rapidly removed into ice-cold artificial cerebrospinal fluid (ACSF), containing the following in mM: NaCl, 125; KCl, 2.5; $MgCl_2$, 1; NaH_2PO4, 1.25; $CaCl_2$, 2; $NaHCO_3$, 26; Dextrose, 25; osmolality 315–320 mOsm, bubbled with 95% O_2/5% CO_2. Acute sagittal slices (300 μm) were obtained with a DTK-1000 Microslicer (Dosaka, Japan) or a VT1000S microtome (Leica Microsystems, Germany) and were moved to 37°C ACSF for 10 min, after which they were allowed to cool to room temperature. All experiments were carried out at 32°C–35°C. Patch pipettes for somatic (5–8 MΩ) and dendritic (15–25 MΩ) whole-cell recordings were pulled on a DMZ Universal Puller (Zeitz, München, Germany) or a PC-10 puller (Narishige, Japan) and filled with (in mM) KCl, 20; K-Gluconate, 100; K-HEPES, 10; MgATP, 4; NaGTP, 0.3; Na-Phosphocreatine, 10; and 0.1% w/v Biocytin, adjusted with KOH to pH 7.2–7.4, and with sucrose to 290–295 mOsm. For 2PLSM imaging (see below), 10–25 μM Alexa Fluor 594 and/or 180 μM Fluo-5F pentapotassium salt (Invitrogen, Carlsbad, CA) were added to the internal solution. No plasticity experiments were performed in the presence of Fluo-5F, to avoid possible Ca^{2+} buffering effects on the induction of plasticity; nor was Fluo-5F ever loaded presynaptically.

Extracellular stimulation electrodes (tip diameters of 2–10 μm) were filled with ACSF and were typically placed in L1 or L2/3 (except for in Figure S3 and a small number of data points in Figure 3, in which case electrode placement was in L5) under direct visual control, approximately 5–50 μm from the apical dendrites of recorded L5 cells. Stimulation electrode position and stimulation strength were adjusted until responses were of the desired amplitude and rise time. In the same manner, polysynaptic responses (of variable latency and/or with multiple peaks) and inhibitory contamination (hyperpolarizing responses and/or dramatic facilitation) were selected against. Collection of baseline responses was only begun once satisfactory responses were obtained (typically 2–10 min). Inhibition was never blocked, to avoid possible effects of tonic GABA blockade on dendritic excitability. In some experiments (Figure 5B), the endocannabinoid CB1 receptor blocker AM251 (Gatley et al., 1996) (Tocris Cookson, UK) was added to the ACSF at 0.9 μM (Sjöström et al., 2003, 2004) to block CB1 receptors.

Whole-cell recordings were obtained using identical BVC-700A amplifiers (Dagan Corporation, Minneapolis, MN) controlled by custom electrophysiology software (Sjöström et al., 2001) running on Igor Pro (WaveMetrics Inc., Lake Oswego, OR). Sweeps were filtered at 10 kHz and acquired using MIO-16E boards (National Instruments, Austin, TX) or ITC-18 boards (InstruTECH, Port Washington, NY) on Macintosh (Apple Computer, Cupertino, CA) or Dell (Dell Computers, Round Rock, TX) computers. Series resistance and capacitance were compensated for dendritic recordings (R_{series} < 50 MΩ).

Pyramidal L2/3 and thick-tufted L5 neurons were identified at 600x magnification (Olympus BX51WI, Olympus, Melville, NY) using infrared Dodt contrast optics (Luigs and Neumann, Ratingen, Germany). We searched for unitary connections using quadruple whole-cell recordings to increase the yield (Figures 1A and 1B), as cortical L2/3-to-L5 and L5-to-L5 connectivity is low (10%–15%) (Song et al., 2005; Thomson and Bannister, 1998). The L2/3-to-L5 connectivity depended steeply on the distance between the L2/3 soma and the L5 apical dendrite (data not shown; see also Thomson and Bannister, 1998), so recorded L2/3 neurons were typically within 50 μm of the nearest L5 apical dendrite. The identity of the neurons was typically verified by morphology, either during the experiment by 2PLSM imaging of Alexa Fluor 594 fluorescence (see below), or after the experiments by biocytin histochemistry (Vectastain ABC Elite kit, Vector Labs, Burlingame, CA). Dendritic recordings were made from L5 neurons at distances up to 450 μm from the soma using standard techniques (Stuart and Häusser, 2001).

Induction of Long-Term Plasticity

After obtaining giga-ohm seals on four neurons, we established the whole-cell configuration in quick succession to prevent unequal LTP dialysis. Connectivity was assessed by averaging 10–40 traces. APs were evoked by 5 ms long current injections (1.0–1.8 nA) in each neuron every 10 s throughout the entire experiment, except during induction (see below). During baseline, APs were displaced by at least 700 ms relative to each other, to avoid accidental induction of long-term plasticity. Experiments were terminated or later discarded if membrane potential changed more than 8 mV or input resistance (measured from 250 ms long 25 pA hyperpolarizing pulses in each sweep) changed more than 30%.

Unless otherwise specified, the plasticity induction protocol consisted of pairing 5 APs at 50 Hz in the pre- and the postsynaptic cells, with pre- leading post- by +10 ms, repeated 15 times every 10 s. The 50 Hz induction frequency was chosen because it yields a maximal amount of timing-dependent LTP in L5-to-L5 pairs (Markram et al., 1997b; Sjöström et al., 2001), so it was likely to highlight proximal-distal differences in plasticity; we note that other induction protocols may result in more or less of a proximal-distal difference. In particular, induction above the critical frequency for distal Ca^{2+} spikes (Larkum et al., 1999a) is likely not to adhere to the learning rule we propose here. After the induction, baseline was resumed for as long as possible. In extracellular stimulation experiments, we employed the same procedure, substituting one extracellular stimulation event for every presynaptic AP evoked in the corresponding paired recording.

Analysis and Statistics

Plasticity experiments shorter than 40 min were not included. Amounts of potentiation and depression were measured as previously described (Sjöström et al., 2001, 2003, 2004): the average response starting 10 min after the induction until the end of the recording was divided by the average response obtained during the initial baseline period. EPSP rise time was measured as the time between 20% and 80% of the peak of the pre-pairing average response.

Comparisons were made using unpaired Student's t test for equal means. Statistically significant change for a particular condition was tested by comparing it with the specified control condition. Statistically significant change in individual experiments was assessed by comparing the post-pairing period to the baseline period. Means are reported as ± SEM. All curve fits were done by least-square gradient descent in Igor Pro. A sigmoidal fit and a sliding t test were used as independent methods to find the threshold for LTP (Figure 4C; also see text). The sliding t test was performed as previously described (Sjöström et al., 2001): The sliding threshold d_t was varied between 0.25 and 3 mV, and a t test was performed for data points above and below d_t. A distinct minimum p value for which $p < 0.001$ suggested a threshold at the corresponding d_t value.

Two-Photon Laser-Scanning Microscopy

All imaging experiments were performed using a custom-built two-photon microscope (Prairie Technologies, Middleton, WI) with a Ti: Sapphire laser (MaiTai, Spectra-Physics) tuned to 800–840 nm. Data was acquired using ScanImage software (Pologruto et al., 2003) running in Matlab (MathWorks, Natick, MA). Ca^{2+} imaging was begun about 1 hr after break-in, to allow for dye wash-in and diffusional equilibration. Ca^{2+} responses were measured by line scans (2.3 ms/line, 256 or 512 lines) of the change in green fluorescence of the Ca^{2+}-sensitive dye Fluo-5F normalized to red Ca^{2+}-insensitive Alexa 594 fluorescence (dG/R). We note that the choice of Fluo-5F is appropriate for examining bAP-induced Ca^{2+} influx in L5 neurons (Yasuda et al., 2004), but is less ideal for relatively small Ca^{2+} influx due to subthreshold activity, so the absence of Ca^{2+} signal due to EPSPs alone does not suggest that there is no Ca^{2+} influx (Figure 7D, cf. Markram and Sakmann, 1994). To account for possible bleaching, we interleaved line scan conditions (e.g., in Figure 7C: first "APs," then "Depol," then "Both") and repeated them a minimum of six times each every 7–10 s. Ca^{2+} traces in Figures 7C and 7D are filtered versions of line scan averages (7 point box filter). The order of line scan positions (Figures 7C–7E) was approximately randomized. The threshold for AP-EPSP-induced supralinear Ca^{2+} signals (see text) was determined just below the first apical dendritic branch point. Ca^{2+} signals were extracted from individual, unfiltered line scan traces, and measured in a 10–25 ms window centered on the peak of the filtered line scan average. Cell morphologies (e.g., Figures 1A and 1B) are pseudocolored maximum intensity projections (3D rendered using OsiriX; Rosset et al., 2004) of anisotropic diffusion-filtered (Broser et al., 2004) red Alexa fluorescence stacks. Dendritic diameters (Figure S2) were measured by Gaussian fits to Alexa fluorescence linescans perpendicular to the dendrite and taking the half-width as the dendritic diameter. This approach was validated in a neuron reconstructed using Neurolucida (see below). Diameters ranged between 0.8–5.0 μm, above the resolution limit of 2PLSM.

Neuronal Reconstructions and Computer Modeling

Putative synaptic contacts were identified by axonal-dendritic crossover points (<1 μm separation), either using 2PLSM imaging of Alexa-filled cells in acute slices (see Figure S1) or using bright-field microscopy with a 100x oil objective (N.A. 1.3, Olympus BX51, Olympus, Melville, NY) of neurons histologically processed for biocytin (see above and Figure 2), or both. Neurons were reconstructed using Neurolucida (MicroBrightField Inc., Magdeburg, Germany) with a 40x objective.

Compartmental modeling was performed using the NEURON simulation environment (Hines and Carnevale, 1997) using a detailed 3D reconstruction of an L5 pyramidal neuron. Passive parameters were $C_m = 1\ \mu F/cm^2$, $R_m = 50\ k\Omega cm^2$, with $V_{rest} = -70$ mV. Spines were accounted for by scaling dendritic membrane capacitance and conductances by a factor of two (Holmes, 1989; Shelton, 1985). Synapses were placed at a range of dendritic locations and consisted of identical AMPA conductances with $E_{rev} = 0$ mV, $g_{max} = 1$ nS, $\tau_{rise} = 0.2$ ms, and $\tau_{decay} = 1.7$ ms (Häusser and Roth, 1997). R_i (182 Ωcm) was adjusted by fitting the model 20%–80% EPSP rise time to the EPSP rise time from the actual paired recording (Figures 2A and 2C), assuming equal contribution from each putative synaptic contact. For the range of anatomically identified synaptic contacts at L2/3-to-L5 connections (Figure 2E), the simulated somatic EPSP amplitudes ranged from 0.16–0.59 mV, which is comparable to the L2/3-to-L5 unitary EPSP amplitudes of actual recordings (0.42 ± 0.10 mV, n = 21). We also made compartmental models from two additional L5 pyramidal cell morphologies using the same approach, and we obtained similar results. Time step for all simulations was 10 μs.

Supplemental Data

The Supplemental Data for this article can be found online at http://www.neuron.org/cgi/content/full/51/2/227/DC1/.

Acknowledgments

We thank Alanna Watt, Arnd Roth, Kazuo Kitamura, Mickey London, Niraj Desai, and Julian Jack for help and useful discussions, and Latha Ramakrishnan for assistance with histology and Neurolucida reconstructions. This work was supported by grants from the Wellcome Trust and the Gatsby Foundation. P.J.S. was supported by a Marie Curie Intra-European Fellowship.

Received: January 12, 2006
Revised: May 2, 2006
Accepted: June 19, 2006
Published: July 19, 2006

References

Artola, A., Bröcher, S., and Singer, W. (1990). Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. Nature *347*, 69–72.

Bekkers, J.M. (2000). Properties of voltage-gated potassium currents in nucleated patches from large layer 5 cortical pyramidal neurons of the rat. J. Physiol. *525*, 593–609.

Bernard, C., and Johnston, D. (2003). Distance-dependent modifiable threshold for action potential back-propagation in hippocampal dendrites. J. Neurophysiol. *90*, 1807–1816.

Broser, P.J., Schulte, R., Lang, S., Roth, A., Helmchen, F., Waters, J., Sakmann, B., and Wittum, G. (2004). Nonlinear anisotropic diffusion filtering of three-dimensional image data from two-photon microscopy. J. Biomed. Opt. *9*, 1253–1264.

Coogan, T.A., and Burkhalter, A. (1990). Conserved patterns of cortico-cortical connections define areal hierarchy in rat visual cortex. Exp. Brain Res. *80*, 49–53.

Cummings, J.A., Mulkey, R.M., Nicoll, R.A., and Malenka, R.C. (1996). Ca^{2+} signaling requirements for long-term depression in the hippocampus. Neuron *16*, 825–833.

Debanne, D., Gähwiler, B.H., and Thompson, S.M. (1996). Cooperative interactions in the induction of long-term potentiation and depression of synaptic excitation between hippocampal CA3–CA1 cell pairs in vitro. Proc. Natl. Acad. Sci. USA *93*, 11225–11230.

Destexhe, A., Rudolph, M., and Pare, D. (2003). The high-conductance state of neocortical neurons in vivo. Nat. Rev. Neurosci. *4*, 739–751.

Froemke, R.C., Poo, M.M., and Dan, Y. (2005). Spike-timing-dependent synaptic plasticity depends on dendritic location. Nature *434*, 221–225.

Gatley, S.J., Gifford, A.N., Volkow, N.D., Lan, R., and Makriyannis, A. (1996). ^{123}I-labeled AM251: a radioiodinated ligand which binds in vivo to mouse brain cannabinoid CB_1 receptors. Eur. J. Pharmacol. *307*, 331–338.

Goldberg, J., Holthoff, K., and Yuste, R. (2002). A problem with Hebb and local spikes. Trends Neurosci. *25*, 433–435.

Golding, N.L., and Spruston, N. (1998). Dendritic sodium spikes are variable triggers of axonal action potentials in hippocampal CA1 pyramidal neurons. Neuron *21*, 1189–1200.

Golding, N.L., Staff, N.P., and Spruston, N. (2002). Dendritic spikes as a mechanism for cooperative long-term potentiation. Nature *418*, 326–331.

Hansel, C., Artola, A., and Singer, W. (1997). Relation between dendritic Ca^{2+} levels and the polarity of synaptic long-term modifications in rat visual cortex neurons. Eur. J. Neurosci. *9*, 2309–2322.

Häusser, M., and Roth, A. (1997). Estimating the time course of the excitatory synaptic conductance in neocortical pyramidal cells using a novel voltage jump method. J. Neurosci. *17*, 7606–7625.

Hebb, D.O. (1949). The Organization of Behavior (New York: Wiley).

Hines, M.L., and Carnevale, N.T. (1997). The NEURON simulation environment. Neural Comput. *9*, 1179–1209.

Hoffman, D.A., and Johnston, D. (1999). Neuromodulation of dendritic action potentials. J. Neurophysiol. *81*, 408–411.

Hoffman, D.A., Magee, J.C., Colbert, C.M., and Johnston, D. (1997). K^+ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons. Nature *387*, 869–875.

Holmes, W.R. (1989). The role of dendritic diameters in maximizing the effectiveness of synaptic inputs. Brain Res. *478*, 127–137.

Katz, L.C., and Shatz, C.J. (1996). Synaptic activity and the construction of cortical circuits. Science *274*, 1133–1138.

Koester, H.J., and Sakmann, B. (1998). Calcium dynamics in single spines during coincident pre- and postsynaptic activity depend on relative timing of back-propagating action potentials and subthreshold excitatory postsynaptic potentials. Proc. Natl. Acad. Sci. USA *95*, 9596–9601.

Korngreen, A., and Sakmann, B. (2000). Voltage-gated K+ channels in layer 5 neocortical pyramidal neurones from young rats: subtypes and gradients. J. Physiol. *525*, 621–639.

Larkum, M.E., Kaiser, K.M.M., and Sakmann, B. (1999a). Calcium electrogenesis in distal apical dendrites of layer 5 pyramidal cells at a critical frequency of back-propagating action potentials. Proc. Natl. Acad. Sci. USA *96*, 14600–14604.

Larkum, M.E., Zhu, J.J., and Sakmann, B. (1999b). A new cellular mechanism for coupling inputs arriving at different cortical layers. Nature *398*, 338–341.

Larkum, M.E., Zhu, J.J., and Sakmann, B. (2001). Dendritic mechanisms underlying the coupling of the dendritic with the axonal action potential initiation zone of adult rat layer 5 pyramidal neurons. J. Physiol. *533*, 447–466.

Levy, W.B., and Steward, O. (1979). Synapses as associative memory elements in the hippocampal formation. Brain Res. *175*, 233–245.

Linden, D.J. (1999). The return of the spike: postsynaptic action potentials and the induction of LTP and LTD. Neuron *22*, 661–666.

Lisman, J., and Spruston, N. (2005). Postsynaptic depolarization requirements for LTP and LTD: a critique of spike timing-dependent plasticity. Nat. Neurosci. *8*, 839–841.

Magee, J.C., and Johnston, D. (1997). A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. Science *275*, 209–213.

Malenka, R.C., and Nicoll, R.A. (1999). Long-term potentiation–a decade of progress? Science *285*, 1870–1874.

Markram, H., and Sakmann, B. (1994). Calcium transients in dendrites of neocortical neurons evoked by single subthreshold excitatory postsynaptic potentials via low-voltage-activated calcium channels. Proc. Natl. Acad. Sci. USA *91*, 5207–5211.

Markram, H., Lübke, J., Frotscher, M., Roth, A., and Sakmann, B. (1997a). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. J. Physiol. *500*, 409–440.

Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997b). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science *275*, 213–215.

McNaughton, B.L., Douglas, R.M., and Goddard, G.V. (1978). Synaptic enhancement in fascia dentata: cooperativity among coactive afferents. Brain Res. *157*, 277–293.

Pologruto, T.A., Sabatini, B.L., and Svoboda, K. (2003). ScanImage: flexible software for operating laser scanning microscopes. Biomed. Eng. Online: http://www.biomedical-engineering-online.com/content/2/1/13.

Rall, W. (1967). Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic input. J. Neurophysiol. *30*, 1138–1168.

Redman, S., and Walmsley, B. (1983). The time course of synaptic potentials evoked in cat spinal motoneurones at identified group Ia synapses. J. Physiol. *343*, 117–133.

Rosset, A., Spadola, L., and Ratib, O. (2004). OsiriX: an open-source software for navigating in multidimensional DICOM images. J. Digit. Imaging *17*, 205–216.

Rumsey, C.C., and Abbott, L.F. (2004). Equalization of synaptic efficacy by activity- and timing-dependent synaptic plasticity. J. Neurophysiol. *91*, 2273–2280.

Schiller, J., Schiller, Y., Stuart, G., and Sakmann, B. (1997). Calcium action potentials restricted to distal apical dendrites of rat neocortical pyramidal neurons. J. Physiol. *505*, 605–616.

Shelton, D.P. (1985). Membrane resistivity estimated for the Purkinje neuron by means of a passive computer model. Neuroscience *14*, 111–131.

Sjöström, P.J., and Nelson, S.B. (2002). Spike timing, calcium signals and synaptic plasticity. Curr. Opin. Neurobiol. *12*, 305–314.

Sjöström, P.J., Turrigiano, G.G., and Nelson, S.B. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron *32*, 1149–1164.

Sjöström, P.J., Turrigiano, G.G., and Nelson, S.B. (2003). Neocortical LTD via coincident activation of presynaptic NMDA and cannabinoid receptors. Neuron *39*, 641–654.

Sjöström, P.J., Turrigiano, G.G., and Nelson, S.B. (2004). Endocannabinoid-dependent neocortical layer-5 LTD in the absence of postsynaptic spiking. J. Neurophysiol. *92*, 3338–3343.

Song, S., Sjöström, P.J., Reigl, M., Nelson, S., and Chklovskii, D.B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol. *3*, e68. 10.1371/journal.pbio.0030068.

Sourdet, V., and Debanne, D. (1999). The role of dendritic filtering in associative long-term synaptic plasticity. Learn. Mem. *6*, 422–447.

Stuart, G., Spruston, N., Sakmann, B., and Häusser, M. (1997). Action potential initiation and backpropagation in neurons of the mammalian CNS. Trends Neurosci. *20*, 125–131.

Stuart, G.J., and Sakmann, B. (1994). Active propagation of somatic action potentials into neocortical pyramidal cell dendrites. Nature *367*, 69–72.

Stuart, G.J., and Häusser, M. (2001). Dendritic coincidence detection of EPSPs and action potentials. Nat. Neurosci. *4*, 63–71.

Thomson, A.M., and Bannister, A.P. (1998). Postsynaptic pyramidal target selection by descending layer III pyramidal axons: dual intracellular recordings and biocytin filling in slices of rat neocortex. Neuroscience *84*, 669–683.

Watanabe, S., Hoffman, D.A., Migliore, M., and Johnston, D. (2002). Dendritic K^+ channels contribute to spike-timing dependent long-term potentiation in hippocampal pyramidal neurons. Proc. Natl. Acad. Sci. USA *99*, 8366–8371.

Waters, J., and Helmchen, F. (2004). Boosting of action potential backpropagation by neocortical network activity in vivo. J. Neurosci. *24*, 11127–11136.

Williams, S.R., and Stuart, G.J. (2000). Backpropagation of physiological spike trains in neocortical pyramidal neurons: implications for temporal coding in dendrites. J. Neurosci. *20*, 8238–8246.

Williams, S.R., and Stuart, G.J. (2002). Dependence of EPSP efficacy on synapse location in neocortical pyramidal neurons. Science *295*, 1907–1910.

Williams, S.R., and Stuart, G.J. (2003). Role of dendritic synapse location in the control of action potential output. Trends Neurosci. *26*, 147–154.

Yasuda, R., Nimchinsky, E.A., Scheuss, V., Pologruto, T.A., Oertner, T.G., Sabatini, B.L., and Svoboda, K. (2004). Imaging calcium concentration dynamics in small neuronal compartments. Sci. STKE, http://stke.sciencemag.org/cgi/content/full/OC_sigtrans; stke.2192004pl5.

Yuste, R., and Denk, W. (1995). Dendritic spines as basic functional units of neuronal integration. Nature *375*, 682–684.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Dendritic Computation

Michael London and Michael Häusser

Wolfson Institute for Biomedical Research and Department of Physiology, University College London, London WC1E 6BT; email: m.london@ucl.ac.uk, m.hausser@ucl.ac.uk

Annu. Rev. Neurosci.
2005. 28:503–32

doi: 10.1146/
annurev.neuro.28.061604.135703

Copyright © 2005 by
Annual Reviews. All rights
reserved

0147-006X/05/0721-
0503$20.00

Key Words

dendrites, coding, synaptic integration, spikes, ion channels

Abstract

One of the central questions in neuroscience is how particular tasks, or computations, are implemented by neural networks to generate behavior. The prevailing view has been that information processing in neural networks results primarily from the properties of synapses and the connectivity of neurons within the network, with the intrinsic excitability of single neurons playing a lesser role. As a consequence, the contribution of single neurons to computation in the brain has long been underestimated. Here we review recent work showing that neuronal dendrites exhibit a range of linear and nonlinear mechanisms that allow them to implement elementary computations. We discuss why these dendritic properties may be essential for the computations performed by the neuron and the network and provide theoretical and experimental examples to support this view.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Contents

INTRODUCTION................. 504
THE DENDRITIC "TOOLKIT"... 505
  Computations in Passive Dendrites 505
  Computations in Active Dendrites. 509
EXAMPLES OF REAL-WORLD DENDRITIC COMPUTATION 516
  Directional Selectivity........... 517
  Coincidence Detection in Auditory Neurons...................... 519
  Temporal Integration Over Long Timescales.................... 520
  Image Processing in Dendrites of Fly Neurons.................. 520
  Looming Sensitive Neurons in the Locust................. 522
  Forward Masking of Cricket Songs 522
  Dendritic Mechanisms and Behavior...................... 524
CHALLENGES FOR THE FUTURE: A WISH LIST FOR DENDRITIC COMPUTATION 524
  For Molecular Biologists: Designer Dendrites........... 524
  For Neurophysiologists: Putting Dendrites Back into the Network...................... 526
  For the Theorist: Proving the Benefits of Dendritic Computation.................. 526
CONCLUSIONS................... 527

INTRODUCTION

Brains compute. This means that they process information, creating abstract representations of physical entities and performing operations on this information in order to execute tasks. One of the main goals of computational neuroscience is to describe these transformations as a sequence of simple elementary steps organized in an algorithmic way. The mechanistic substrate for these computations has long been debated. Traditionally, relatively simple computational properties have been attributed to the individual neuron, with the complex computations that are the hallmark of brains being performed by the network of these simple elements. In this framework, the neuron (often called a "Perceptron," "Spin," or "Unit") sums up the synaptic input and, by comparing this sum against a threshold, "decides" whether to initiate an action potential. In computational terms this process includes only one nonlinear term (thresholding), which is usually counted as a single operation. Thus, the neuron operates as a device where analog computations are at some decision point transformed into a digital output signal. Such a design forms the backbone of many artificial neuronal networks, starting from the original work of McCullough & Pitts (1943) to the present day. In this review we argue that this model is oversimplified in view of the properties of real neurons and the computations they perform. Rather, additional linear and nonlinear mechanisms in the dendritic tree are likely to serve as computational building blocks, which combined together play a key role in the overall computation performed by the neuron.

In the first part of this review we describe the dendritic computational toolkit, i.e., the biophysical mechanisms in dendrites, which endow them with potential computational powers. We focus on the most recent findings and group them according to the type of computations being carried out. In the second part of the review, we present several examples where the role of dendritic mechanisms in computation has strong circumstantial support, such as directional selectivity in retinal neurons or coincidence detection in the auditory system. We conclude with a discussion of how we can ultimately prove the role of dendrites in neural computation and outline a process for achieving this goal. Throughout the review, we focus on the *online* aspects of computation. Naturally, the results of such computations must ultimately be read out and stored within the network. Given that dendrites are also the site where synaptic plasticity takes place, their properties are likely to affect the induction and expression of plasticity. These issues have been discussed in several recent reviews (Häusser & Mel 2003, Linden 1999, Mainen 1999, Chklovskii et al. 2004), and we therefore do not address this aspect in detail here.

THE DENDRITIC "TOOLKIT"

Why discuss what a neuron does in terms of computation? We use a simple example to illustrate the problem. When one watches a movie, each frame is presented for about 50 ms, during which time we have to process any changes from the last frame before the next frame appears. To a first approximation the timescale of the computational cycle of a neuron involved in this computation is thus on the order of milliseconds to tens of milliseconds. Because a typical neuron in the brain receives thousands of synaptic inputs, a neuron involved in processing the visual input thus receives hundreds to thousands of 50-ms input spike trains. But the neuron has only one axon, which provides the output signal, and thus the final conversion represents a compression into a much smaller amount of information. Because there are so many inputs, an essential feature of this transformation is the amount of input information that must be discarded by the neuron. This process is analogous to what some mathematical functions are doing: projecting a huge space onto a narrow one. Thus, the neuron throws away the irrelevant information and selects only the information relevant to its function. The essence of computing a function in computer science is usually considered to be implementing an algorithm, that is, a sequence of simple computational steps organized in a flowchart that leads from the input to the output (**Figure 1**). Are there ways to deconstruct into such simpler building blocks the very complex mapping done by a neuron? Can we identify the crucial steps that occur during this process? Are decisions taken only at the final step of transforming the somatic voltage into an action potential, or are there decision points on the way?

Although the flowchart representation may not be the most appropriate and comprehensive way to describe what neurons are doing, it can nevertheless be very instructive to define and explore the computational building blocks (i.e., the boxes on the flowchart) that neurons can perform. During the past decade a rapidly increasing number of studies investigating signaling mechanisms in dendrites have appeared, and several recent reviews have discussed these findings in depth (Euler & Denk 2001, Häusser et al. 2000, Koch & Segev 2000, Magee 2000, Segev & London 2000, Williams & Stuart 2003). Rather than listing all the known facts about dendritic signaling, in this section we focus on identifying the unique dendritic mechanisms that can act as computational building blocks. We start with a brief discussion of how passive dendrites transform their inputs and implement nonlinear interactions between synaptic inputs. Then we review the role of dendritic voltage-dependent channels and conclude by evaluating the possible role of dendrites in chemical computation.

Computations in Passive Dendrites

It is important to recognize that the passive properties of the dendritic tree provide the backbone for the electrical signaling in dendrites, even when they are richly endowed with voltage-dependent ionic currents. For example, the threshold for initiation of a dendritic spike is determined in part by the availability of sodium channels, but perhaps even more by the passive load of the surrounding dendrites, which dictates how much of the input current will depolarize the membrane and how much will flow axially toward other dendritic regions (Segev & London 1999). Thus, understanding the passive properties of dendritic trees remains crucial for understanding computation in dendritic trees. Aside from their role in regulating the conditions for active dendritic signaling, the passive properties

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

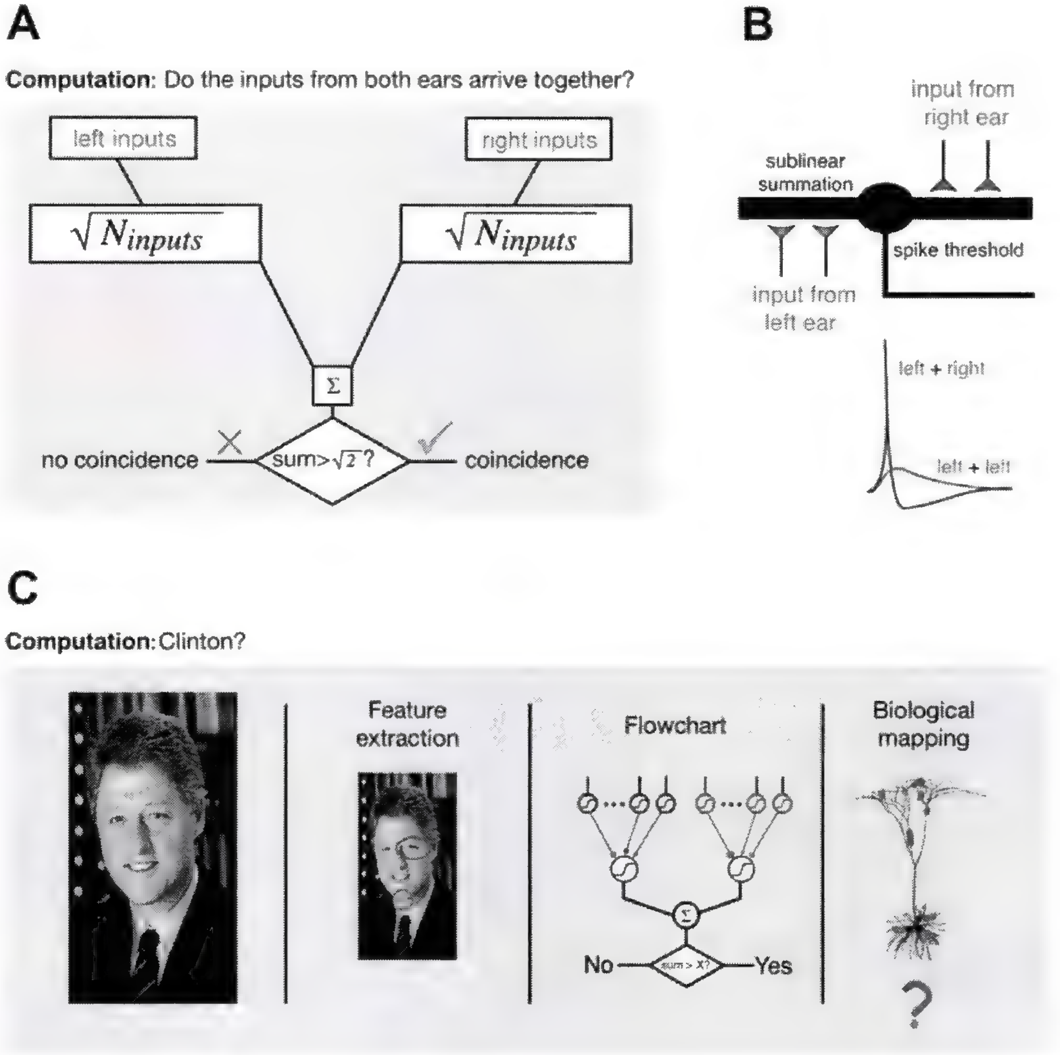

A
Computation: Do the inputs from both ears arrive together?
left inputs
right inputs
√Ninputs
√Ninputs
Σ
no coincidence
sum>√2 ?
coincidence
B
input from right ear
sublinear summation
spike threshold
input from left ear
left + right
left + left
C
Computation: Clinton?
Feature extraction
Flowchart
Biological mapping
Σ
No
sum > X?
Yes
?

Figure 1

Dendritic computation. The task of a brainstem auditory neuron performing coincidence detection in the sound localization system of birds is to respond only if the inputs arriving from both ears coincide in a precise manner (10–100 μs), while avoiding a response when the input comes from only one ear. (*A*) A flowchart of a simplistic algorithm to achieve this computation. Inputs to each ear are summed sublinearly, and the input from both ears is then summed linearly and compared with threshold. Thus only if there are inputs from both ears will the sum exceed the threshold, whereas if the input arrives only to one ear the output is not large enough. (*B*) Agmon-Snir et al. (1998) showed that dendrites of these neurons might implement a similar algorithm. Inputs from each ear arrive on one dendrite. Sublinear summation is achieved by the mutual shunting of the excitatory inputs, and the threshold is implemented via the spike-generation mechanism. (*C*) Neurons are known to be involved in much more sophisticated computations, such as face recognition (Kreiman et al. 2000). An algorithm to solve a face recognition task is one of the holy grails of computer science. At present, we do not know precisely how single neurons are involved in this computation. An essential first step is feature extraction from the image, which clearly involves a lot of network preprocessing before features are fed into the individual cortical neuron. The flowchart implements a three-layer model of dendritic processing (see Häusser & Mel 2003) to integrate the input. The way such a flowchart is mapped onto the real geometry of a cortical pyramidal neuron (*right panel*) remains unknown.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

of dendrites provide computational functions in their own right, as discussed below.

Delay lines via dendritic filtering. In terms of signal propagation, dendrites behave like electrical cables with medium-quality insulation. As such, passive dendrites linearly filter the input signal as it spreads to the site of initiation, where it is compared with the threshold. This filtering tends to attenuate the dendritic signal as a function of the distance it travels and the frequency of the original signal. Thus a brief and sharp excitatory postsynaptic potential (EPSP) that originates in the dendrite will be transformed into a much smaller and broader signal when it arrives at the soma (**Figure 2**). As a consequence, the time-to-peak of a synaptic event, and thus the delay of the resulting output spikes, depends on the location of the synapse in the dendritic tree. Rall (1964) recognized that this property may be exploited to perform simple computations. First, for single inputs, by acting as a delay line, the dendrites can thus "label" particular inputs on distinct regions of the dendritic tree by the latency of the resulting output spikes. In fact, EPSPs with different somatic shape are likely to affect the somatic output spike trains in different ways (Fetz & Gustafsson 1983). Second, for multiple inputs, the time course of the somatic voltage response depends on the temporal order of activation of the dendritic synapses (in contrast to the scenario where they are all located on the soma) (Rall 1964; see below, and see also **Figure 6**).

Parallel processing and local computations. Synaptic inputs onto dendrites do not only inject current but also locally change the membrane conductance to certain ions. This leads to a nonlinear interaction between multiple inputs if they colocalize in time and space. When two excitatory inputs are active together at short distance, each depolarizes the membrane and reduces the driving force for the other input, and thus, theoretically, the response to the simultaneous activation is smaller than the sum of the individual responses (Rall et al. 1967). In this context, dendrites might be beneficial because they enable the spatial separation of inputs to minimize their interaction. In some cases, however, this possible sublinear summation may actually be advantageous (Agmon-Snir et al. 1998, section on coincidence detection in auditory neurons, p. 519) (see **Figure 1**). It also provides a mechanism for saturation of inputs, thus preventing overexcitation of the neuron by a group of synapses.

Nonlinear interactions are especially prominent between excitatory synapses and shunting inhibition. Shunting inhibition usually describes inhibition that changes the total conductance of the membrane but does not cause any voltage change when activated on its own. In this case it is more convenient to think of the inhibition as reducing the input resistance of the cell, effectively reducing the voltage response to excitatory current. This property of inhibition can be mathematically formulated as having a divisive effect on excitatory signals (Fatt & Katz 1953), providing a mechanism for simple arithmetic computation in dendrites (Blomfield 1974). Rall was the first to recognize that the effectiveness of this interaction has a strong spatial component. When excitatory and inhibitory inputs are widely separated from each other on different dendritic branches, then the inputs will tend to sum linearly at the soma. In contrast, when the excitatory and inhibitory inputs are located adjacent to each other, then the inhibition can produce a highly nonlinear "shunting" of the excitatory input (Rall 1964). An elegant recent experimental study by Liu (2004) demonstrated that such an inhibitory effect can be localized to a single dendritic branch. Theoretical work has shown that inhibition is also effective when it is located on the path between the excitatory input and the soma (Jack et al. 1975, Koch et al. 1983, Rall 1964). Thus, the relative location of synaptic inhibition versus excitation determines whether

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

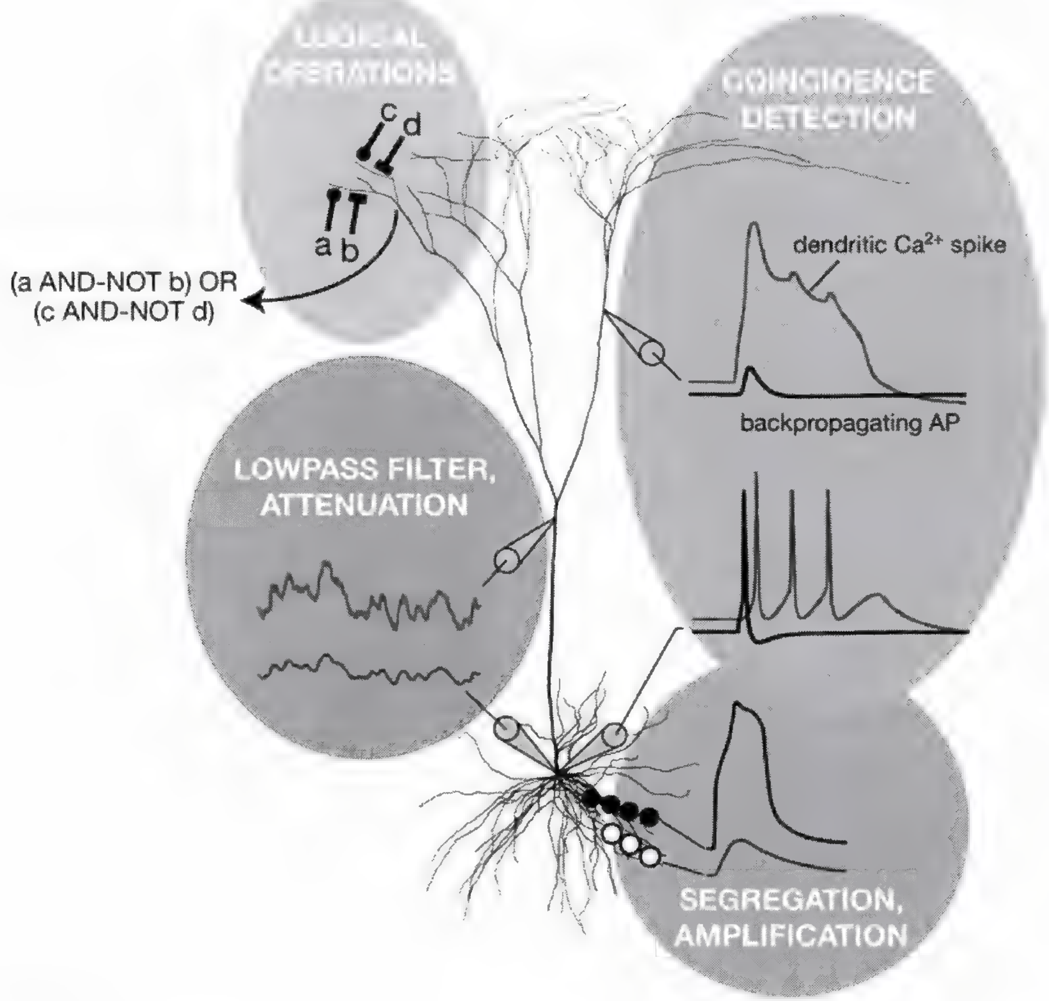

LOGICAL OPERATIONS
c
d
a
b
(a AND-NOT b) OR (c AND-NOT d)
COINCIDENCE DETECTION
dendritic Ca^{2+} spike
backpropagating AP
LOWPASS FILTER, ATTENUATION
SEGREGATION, AMPLIFICATION

Figure 2

The dendritic computational toolkit. A schematic figure highlighting four key dendritic mechanisms, mapped onto a layer 5 pyramidal neuron morphology, which can allow dendrites to act as computational elements. These mechanisms can coexist in the same neuron and be active in parallel or in a hierarchical manner. *Bottom left*: Passive dendrites act as passive filters. A high-frequency fluctuating current injected in the dendritic pipette will evoke high-frequency and large-amplitude local voltage responses, but the response recorded by the somatic pipette will be attenuated and smoothed (low pass filtered). *Top left*: Nonlinear interaction between excitation and shunting inhibition on small dendritic branches can implement logical operations. The branch marked by an arrow sums up the current from the two subtrees, such that its output would be a logical OR on their output. Each of the subtrees will in turn inject current if and only if the excitation AND-NOT the inhibition will be active. *Bottom right:* Dendrites can help reduce or amplify the mutual interaction between excitatory inputs. Excitatory inputs to the same branch tend to sum sublinearly, whereas inputs on different branches sum linearly. Thus mapping input to different branches can reduce this effect. In neurons with active dendrites, however, clusters of inputs active synchronously on the same branch can evoke a local dendritic spike, which leads to significant amplification of the input. Synapses onto a different branch (*open circles*) are only slightly influenced by this spike. *Top right*: In layer 5 cortical pyramidal neurons, as depicted here, coincidence detection between the apical and basal dendritic compartments is achieved by active dendritic mechanisms. A backpropagating action potential, which coincides with a distal synaptic input, will trigger a dendritic Ca^{2+} spike, which depolarizes the whole apical dendrite and drives a burst of spikes in the axon. See text for further details.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

inhibition predominantly counteracts a specific set of (neighboring) excitatory synapses or whether it acts on the global set of excitatory synapses.

Although inhibition can act in a graded manner, it has been predicted that, in principle, synaptic inhibition may be able to veto an excitatory signal effectively, depending on the location and the strength of the inhibitory conductances (Jack et al. 1975, Koch et al. 1983, Rall 1964). For example, the result of the combined operation of a neighboring pair of excitatory and inhibitory inputs will cause somatic depolarization if and only if excitatory input AND NOT inhibitory input is active (**Figure 2**). This AND-NOT function is a Boolean logical operation, of the exact same kind implemented in modern computers and studied in mathematical computational theory. Whether dendrites really implement a network of Boolean gates is not clear, but this is exactly the kind of formalism required for a deep understanding of dendritic computation, namely a formal mathematical entity that would clarify the operations performed by dendrites. Koch et al. (1983) cleverly showed that logic operations can be linked to the less formal notions of computation used by physiologists to devise a model of a retinal ganglion neuron that has a directional selectivity to moving visual inputs (**Figures 2** and **6**; see below). Because the branch points in the dendritic tree can be seen as summing up the currents in individual branches, each tree can be seen as summing over many logical gates, and thus the whole dendrite can implement complex functions. Note that a key issue in the implementation of such a mechanism is the addressing of the right synapses to the right dendrite. In fact, as a general rule, as we see below, any computation that exploits local nonlinearity mechanisms is bound to require the addressing of the relevant synaptic inputs to the relevant locality in the dendrite (Mehta 2004, Poirazi & Mel 2001). It remains to be seen whether the power of dendritic computation can itself provide a constraint for targeting of synaptic inputs at the appropriate locations and whether such addressing indeed takes place as a basic phenomenon in the brain (Chklovskii et al. 2004).

Computations in Active Dendrites

Dendritic excitability as a feedback mechanism. Solely on the basis of anatomical observations, Cajal formulated the law of dynamic polarization (Cajal 1911), which states that in the nervous system information flows in one direction: from dendrites to soma to axon. In the past decade it has become clear that in many types of neurons the presence of excitable ionic currents in the dendrites supports dendritic action potentials that travel in the reverse direction, from the soma into the dendrites (Stuart et al. 1997). Computationally this "backpropagation" has major consequences because it implies that the neuron is no longer an open-loop system but has an internal mechanism of feedback. It is thus no longer the case that feedback is a property only of the network, but rather it is a salient property of each element of the network. Moreover, the feedback conveyed by the backpropagating action potential is highly sophisticated and has many important consequences for dendritic function, and also for synaptic plasticity (Magee & Johnston 1997, Linden 1999). For example, a single backpropagating action potential can activate slow dendritic voltage-gated ionic currents, which in turn flow back towards the spike initiation zone in the axon, often resulting in additional action potentials (see below). Thus, the somatic action potential can under favorable conditions trigger a burst due to its interaction with the dendrites (Carruth & Magee 1999; Williams & Stuart 1999). Modeling studies show that this interaction between soma and dendrites can be captured by a reduced, two-compartment model of the neuron and critically depends on the coupling coefficient between the two compartments. This coupling is governed by dendritic morphology and the distribution and properties of dendritic

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

voltage-gated channels and synaptic activity (Doiron et al. 2002, Mainen & Sejnowski 1996, Pinsky & Rinzel 1994, Schaefer et al. 2003, Vetter et al. 2001). The firing patterns of neurons are thus potentially tunable simply by changing dendritic properties. The interplay between somatic spikes and dendritic response could be exploited computationally, e.g., as a slope detector (Kepecs et al. 2002), or for feature detection in sensory systems (Oswald et al. 2004).

Amplification of synaptic inputs. The fact that passive dendrites attenuate the synaptic input on its way to the soma raises a longstanding question: Why have so many distal synapses if their activity is not going to affect the output whatsoever? Investigators have proposed that other mechanisms are involved in synaptic integration, effectively endowing each synapse with equal "vote" and creating a "dendritic democracy." Resolving the importance of these different mechanisms in different neuronal types is important because the presence or absence of compensatory mechanisms leads to fundamentally different views of neuronal function (Häusser & Mel 2003). Here we briefly outline the various scenarios and the supporting experimental evidence. Four major mechanisms have been proposed: synaptic scaling, synaptic boosting, local dendritic spikes, and global dendritic spikes.

1. Synaptic scaling: In this scenario, the conductances of distal synapses are scaled according to their distance from the soma, so as to equalize their efficacies. First, indirect evidence for this mechanism of "dendritic democracy" was found in motoneurons (Iansek & Redman 1973) and has been more recently supported by studies in hippocampal CA1 pyramidal neurons (Magee 2000). However, this mechanism may not be general because other major types of neurons do not seem to follow this rule (Williams & Stuart 2002).
2. Subthreshold boosting: Inward voltage-dependent dendritic currents can amplify synaptic inputs on their way to the soma, thus compensating for their attenuation. Although it is clear that there exist dendritic currents to support this scenario (Cook & Johnston 1997, 1999; Migliore & Shepherd 2002), there is contradictory experimental evidence regarding whether such boosting plays an important role, and whether it stems from dendritic or somatic currents (Oviedo & Reyes 2002, Schwindt & Crill 1995, Stuart & Sakmann 1995).
3. Local dendritic spikes: A powerful mechanism for overcoming dendritic attenuation is the local dendritic spikes triggered by coactivation of synaptic inputs. The regenerative inward currents required for triggering such spikes can be provided by voltage-gated sodium channels, voltage-gated calcium channels, or synaptically activated N-methyl-D-aspartate (NMDA) receptor channels. Such spikes could be triggered by synaptic inputs or local dendritic current injections. The spatial extent of these dendritic spikes is highly variable, and so is their effect on the somatic voltage (Gasparini et al. 2004, Golding & Spruston 1998, Mel 1993, Polsky et al. 2004, Schiller et al. 1997, Softky 1994, Stuart et al. 1997, Williams & Stuart 2002). Theoretical studies have recently shown that such a mechanism could lead to a substantial increase in the computational power of the neuron (see below; Poirazi & Mel 2001; **Figure 5**).
4. Global dendritic spikes: Layer 5 cortical pyramidal neurons represent an extreme case where, owing to the length of the apical dendrite, many synapses are located so distally that, in the absence of any compensation mechanisms, they would have virtually no effect on the somatic voltage (Cauller & Connors 1994; Stuart & Spruston 1998). Not

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

only is there marked passive attenuation in these neurons, but active currents may further attenuate the signal (see below). Recently experiments have shown that these neurons exhibit a second spike-initiation zone near the main branch point of the apical dendrite (about 500–650 μm from the soma). Investigators have suggested that the distal apical dendritic tree can act as a separate synaptic integration region, having its own separate spike-initiation zone at this location. When the distal compartment crosses the threshold, a dendritic Ca^{2+} spike is initiated, resulting in a huge dendritic depolarization that drives the somatic region to initiate action potentials. In this way the distal compartment communicates with the soma (Larkum et al. 1999b, 2001; Schiller et al. 1997; Williams 2004; Yuste et al. 1994) (**Figure 4**). To some extent this mechanism is also likely to be operational in other types of pyramidal neuron, such as those in hippocampal CA1 (Golding et al. 2002).

Compressive dendritic nonlinearities. Dendrites express not only voltage-gated inward currents, but also they are rich in other classes of voltage-gated currents that counteract regenerative excitation and thus can be thought to be responsible for maintaining the balance of excitability in the dendritic tree. Some of these currents, such as A-type K^+ currents or hyperpolarization-activated inward I_h currents, are located at higher densities in the distal part of the dendrites (Hoffman et al. 1997, Lörincz et al. 2002, Magee 1998, Williams & Stuart 2000; reviewed by Migliore & Shepherd 2002). This arrangement of nonregenerative currents in the dendrites is puzzling because it enhances the attenuation experienced by synaptic inputs. However, in view of the rich complement of excitable currents in the dendrites such mechanisms are likely required to maintain dendritic stability. In addition to their global balancing effect, these currents can take part in more local interactions. One example for such an interaction has been shown for A-type K^+ currents. When a synaptic input is active on a dendritic branch, the depolarization of the branch inactivates A-type K^+ currents in this branch. This in turn facilitates the ability of backpropagating action potentials to invade this branch more easily, which may provide a gating mechanism for plasticity in dendritic branches (Hoffman et al. 1997, Magee et al. 1998).

Coincidence detection. The simplest nonlinear operation is multiplication. In case of binary variables, multiplication is identical to the logical AND operation (the result is 1 if and only if the two inputs are 1), which can also be described in terms of coincidence detection. We have already described how the interaction of excitation and inhibition could implement this operation, but the regenerative inward currents expressed in the dendritic membrane provide numerous alternative mechanisms. These forms of coincidence detection can operate on a highly local scale (down to the level of a few spines) or on the scale of the entire neuron.

Numerous experimental studies have provided evidence that active dendrites can generate local dendritic spikes given synaptic input that is sufficiently clustered in space and time (the latter being the essential requirement for a coincidence detector). Such spikes can be generated by any combination of the voltage-gated regenerative inward currents known to be present in the dendritic membrane. For example, the current driven by NMDA receptor activation is known to be highly voltage dependent. Recently investigators (Cai et al. 2004, Polsky et al. 2004, Schiller et al. 2000) (see **Figure 5**) showed that synchronous synaptic inputs onto the same dendritic branch of layer 5 pyramidal neurons depolarize the membrane and create a positive feedback loop such that the current through the NMDA receptor depolarizes the membrane and recruits more NMDA-mediated current (supported by activation of dendritic Na^+ and Ca^{2+} channels). This all-or-none phenomenon is termed an NMDA spike, and its spatial extent is limited to a short region of the dendrite by active and passive mechanisms. In CA1 pyramidal neurons, a similar local coincidence-detection mechanism exists, based primarily on a different voltage-dependent current (Na^+) (Ariav et al. 2003). These dendritic mechanisms provide the neuron with the ability to detect coincidences in neighboring synaptic inputs on a very fast timescale, previously thought to be restricted to auditory neurons. The same mechanism was previously suggested in models to explain the variability in output spike trains of cortical neurons (Softky 1994).

Dendritic mechanisms also exist for reporting coincident pre- and postsynaptic activity. At distal synapses on the apical dendrite of layer 5 pyramidal neurons, pairing postsynaptic action potentials and synaptic input can trigger highly nonlinear amplification of backpropagating dendritic action potentials via recruitment of voltage-gated Na^+ channels (Stuart & Häusser 2001) (**Figure 3**). A similar supralinear interaction has also been observed in CA1 pyramidal neurons when pairing backpropagating spikes and EPSPs, where the contribution of A-type K^+ channel inactivation is more prominent (Johnston et al. 2000, Magee & Johnston 1997). This form of coincidence detection exhibits a narrow time window (~10 ms), similar to that required for induction of synaptic plasticity when pairing APs and EPSPs, and may thus act as a substrate for the induction of synaptic plasticity.

Finally, in layer 5 pyramidal neurons excitatory synaptic input to the distal apical tuft that coincides with backpropagation of the action potential results in a large dendritic Ca^+ spike, which in turn propagates toward the soma and drives the axon to fire a burst of action potentials (Larkum et al. 1999b, Schiller et al. 1997, Stuart & Häusser 2001) (**Figure 4**). This mechanism thus enables the detection of coincident activation of synaptic inputs to the two major compartments of the dendritic tree, and may thus be involved in reporting simultaneous activity across different cortical layers. This coincidence detection mechanism is potentially tunable, either by changing dendritic geometry or by modulating channel densities and properties (Vetter et al. 2001, Schaefer et al. 2003).

Dendritic subunits: neurons within a neuron. Nonlinear mechanisms in dendrites can vary widely in the spatial extent of the resulting electrical and chemical signals. Some events spread across the entire dendritic tree, whereas others remain very local (Häusser & Mel 2003). Focusing on the local mechanisms, Mel and colleagues (Mel 1993; Poirazi et al. 2003a; Poirazi & Mel 2001; Polsky et al. 2004) have developed a framework that breaks the dendrites into many tiny computational units. The basic nonlinearity in individual branchlets is based on the NMDA spike (Schiller et al. 2000, Schiller & Schiller 2001) and is modeled as a sigmoidal function (**Figure 5**). Each subunit thus integrates its inputs and passes them through a sigmodial nonlinearity function. This gives each piece of dendrite the computational power of a small unit similar to those conventionally used in neural networks. The output of each subunit is conveyed to the soma, in terms of passive dendritic integration. The picture that emerges from this analysis is of a two-layered neuronal network that resides within a single neuron. This analysis is supported by detailed modeling of single neurons, showing that the predictive power of the two-layer neural network description is very good (Poirazi et al. 2003b), and by experiments in layer 5 pyramidal neurons (Polsky et al. 2004). The attractiveness of this approach stems from the fact that two-layer neural networks are general-purpose computation machines, which have been extensively studied and can implement very powerful computations.

This analysis also poses important questions about the way the neuron learns to

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

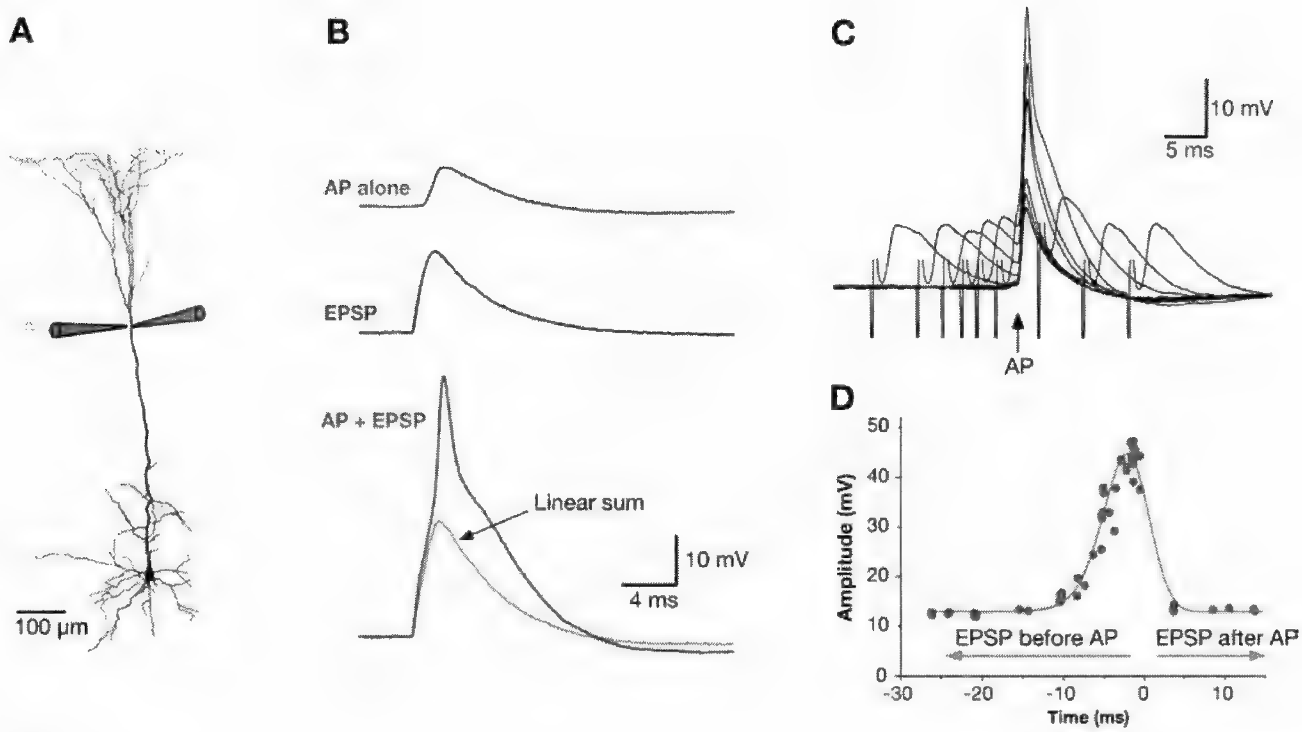


Figure 3

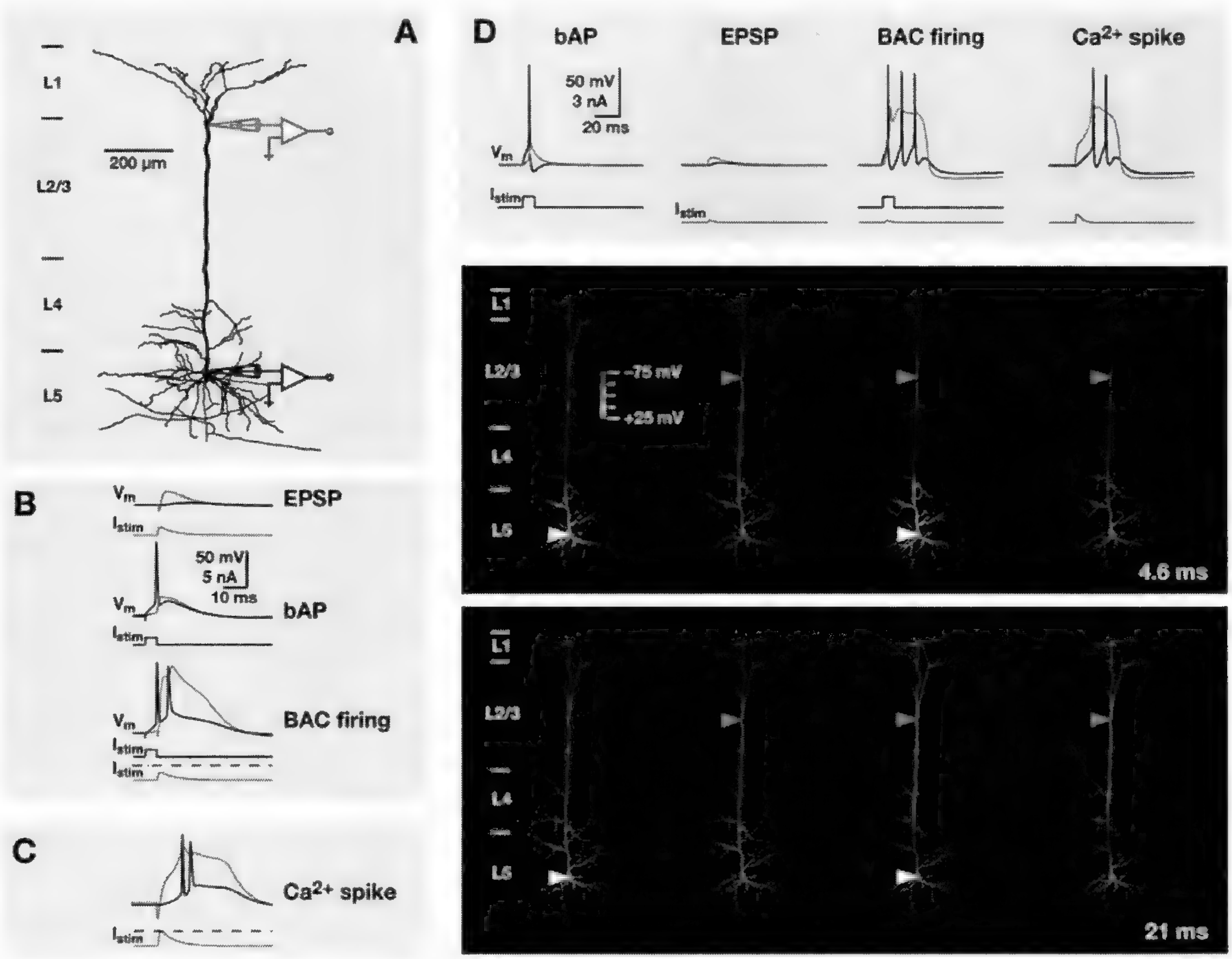
Coincidence detection of EPSPs and action potentials. (*A*) Schematic illustration of the recording configuration. A dendritic recording (*red pipette*) is made from the distal apical dendrite of a layer 5 pyramidal neuron, and synaptic input local to the recording site is activated with a stimulation electrode (*blue pipette*). (*B*) *Top trace*: backpropagating action potential initiated by somatic current injection (2.5 nA) recorded 720 μm from the soma. *Middle trace*: evoked EPSP recorded at the same dendritic location. *Bottom trace*: response to pairing the backpropagating action potential with the EPSP ("AP + EPSP"). For comparison, the linear sum of the action potential plus EPSP is also shown (*grey trace*). Stimulus artifacts are blanked for clarity. (*C*) Superimposed sweeps of evoked EPSPs at different times before and after action potentials initiated by somatic current injections (1.5 nA) at the time indicated by the arrow. Dendritic recording 480 μm from the soma. (*D*) Plot of action potential amplitude (measured at the time of action potential peak) versus the time difference between EPSP and somatic action potential onset (same cell as in *C*). The time of somatic action potential peak is defined as zero; negative time corresponds to when EPSPs were evoked before action potentials, and positive time corresponds to when EPSPs were evoked after action potentials. The smooth line is the fit with a skewed Gaussian. Modified from Stuart & Häusser (2001).

compute its input-output function. In this review we must assume that neurons are learning to compute what they compute. The leading theory for how this is achieved is Hebbian plasticity (Hebb 1949). But if dendrites are implementing these nonlinear subunits, Hebbian plasticity will not exploit the power of this model. Plasticity will help to drive learning within each individual subunit, but the number of inputs in each of these units is small and it is not clear how the relevant inputs will get there in the first place. Poirazi & Mel (2001) have proposed a learning algorithm by which synaptic connections are continuously remodelled until they hit the correct dendritic subunit. They show that such a learning algorithm implemented in a dendritic tree can be much more powerful than the Hebbian learning scheme. Although there exists conflicting evidence for ongoing remodeling of synaptic inputs in the adult brain (Grutzendler et al. 2002, Holtmaat et al. 2005, Mizrahi & Katz 2003, Trachtenberg et al. 2002), it is still an open question whether this algorithm indeed

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

is implemented in real neural circuits. Other challenges for the model are its integration with the dendritic tree. All the model is currently taking from the dendritic tree is the ability to have independent subunits. However, it ignores global nonlinearities and modulations evident in the dendrites. Moreover, in a two-layer neural network, the coefficients from each unit in the first layer to the next are modifiable. Here, in comparison, they are fixed and determined by the dendritic tree. Finally, although the two-layer neural network model reliably predicts the steady-state firing rates of the pyramidal neurons, it neglects the temporal properties of spike firing that have been closely linked to dendritic excitability (e.g., Ariav et al. 2003, Larkum et al. 1999b).

Chemical computation. The expression of voltage-gated calcium channels in the dendritic membrane (Migliore & Shepherd 2002) immediately provides a biochemical readout of electrical excitability. In particular, dendritic calcium signals activated by backpropagating action potentials reliably encode the level of axonal spike firing in apical dendrites of pyramidal cells. This therefore provides a "frequency code" where firing rate is read out using a dendritic biochemical signal (Helmchen et al. 1996). This readout can also have a nonlinear frequency dependence if it involves activation of a dendritic calcium spike (Larkum et al. 1999a). The calcium signal can in turn activate voltage-gated potassium currents, thus acting as a feedback regulator of excitability, which changes the



input-output gain of the neuron (Sobel & Tank 1994) (see below). A similar readout of dendritic excitability can be provided by intracellular dendritic Na^+ signals (Rose & Konnerth 2001), which may in turn regulate excitability via activation of Na^+-activated K^+ channels.

Whereas the various regenerative dendritic mechanisms for coincidence detection discussed above will also be read out via voltage-gated calcium channels to generate large dendritic calcium signals, the biochemical intracellular signaling pathways in dendrites may themselves contribute to coincidence detection in unique ways. For example, the IP3 receptor is cooperatively activated by both calcium and IP3, which allows for coincidence detection of calcium and IP3 delivered by different sources, such as action potentials and synaptic mGluR activation (Nakamura et al. 1999, Wang et al. 2000). This form of coincidence detection can be spatially segregated to particular types of dendrite (Nakamura et al. 2002). A further mechanism for coincidence detection and/or intracellular calcium amplification can be implemented via dendritic calcium-induced calcium release from stores (Emptage et al. 1999). Finally, the lowly calcium-buffering proteins localized in dendrites may themselves allow for a simple form of coincidence detection, generating a supralinear dendritic calcium signal by buffer saturation (Maeda et al. 1999). The large calcium signals generated by these different forms of coincidence detection can remain highly localized (Wang et al. 2000), or they can spread to other regions of the dendritic tree, assisted by further regenerative calcium release from stores (Larkum et al. 2003, Nakamura et al. 2002). Barlow proposed that such processing by intracellular networks can implement a second layer of computation

Figure 4

Coincidence detection across dendritic compartments. (*A*) Reconstruction of a layer 5 pyramidal neuron; the locations of recording pipettes (soma, *black*; dendrite, *red*) are depicted schematically. (*B*) Distal current injection of 1.1 nA in the shape of an EPSP (I_{stim}, *red*) evoked only weak somatic (*black*) depolarization (*upper panel*). Threshold current injection (5 ms) into the soma (*black*) produced an AP that propagated back into the apical dendritic arbor (backpropagating action potential, bAP, *red trace*, *middle panel*). Combination of somatic and dendritic current injection generates several somatic APs and a dendritic Ca^{2+} spike (backpropagating action potential–activated Ca^{2+} spike firing, BAC firing; *lower panel*). The dashed line indicates the current threshold for a dendritic Ca^{2+} spike alone. (*C*) A dendritic Ca^{2+} spike was evoked by 2 nA current injection into the apical dendrite alone. Thus, the bAP reduced the threshold for dendritic Ca^{2+} spike by 0.9 nA (45% coupling). (*D*) A model of channel density distributions and kinetics was constructed to reproduce BAC firing in reconstructed model neurons (Schaefer et al. 2003). The electrical response of the reconstructed model neurons to dendritic and somatic current injection was investigated using the same protocols as in the experiment (*A–C*). (*Upper panels*) bAP: Threshold somatic current injection evoked a bAP ($I_{stim} = 1.9$ nA). EPSP: Distal EPSP-like current injection was adjusted to BAC firing threshold, which was 0.6 nA. Only a small somatic depolarization can be detected ($\Delta V \leq 2.5$ mV). BAC firing: Pairing the bAP with the dendritic EPSP-like current injection resulted in a large and long-lasting dendritic depolarization. Ca^{2+} spike: Large distal EPSP-like current injection (1.7 nA) elicited a Ca^{2+} spike. Thus, the bAP reduced the threshold for dendritic Ca^{2+} spikes by 1.1 nA, which resulted in a coupling of 1.1 nA/1.7 nA = 65%. Voltages were measured at the positions indicated by triangles in lower panels; (*red*: dendritic recording/current injection; *black*: somatic recording/current injection). (*Lower panels*) Same as upper panels but showing membrane potential in the entire dendritic tree. Voltages are color coded as indicated in the upper left. The position of current injection is indicated by the red (dendritic) and white (somatic) arrowheads. At the time of AP initiation (4.6 ms after the beginning of the somatic current injection), depolarization due to the bAP has already spread into the apical dendrite (in the case of bAP and BAC firing). After 21 ms, the voltage deflection due to the bAP has decayed back to baseline. Note that the spread of depolarization is almost the same for a dendritically elicited Ca^{2+} spike and BAC firing. Modified from Larkum et al. (1999b) and Schaefer et al. (2003).

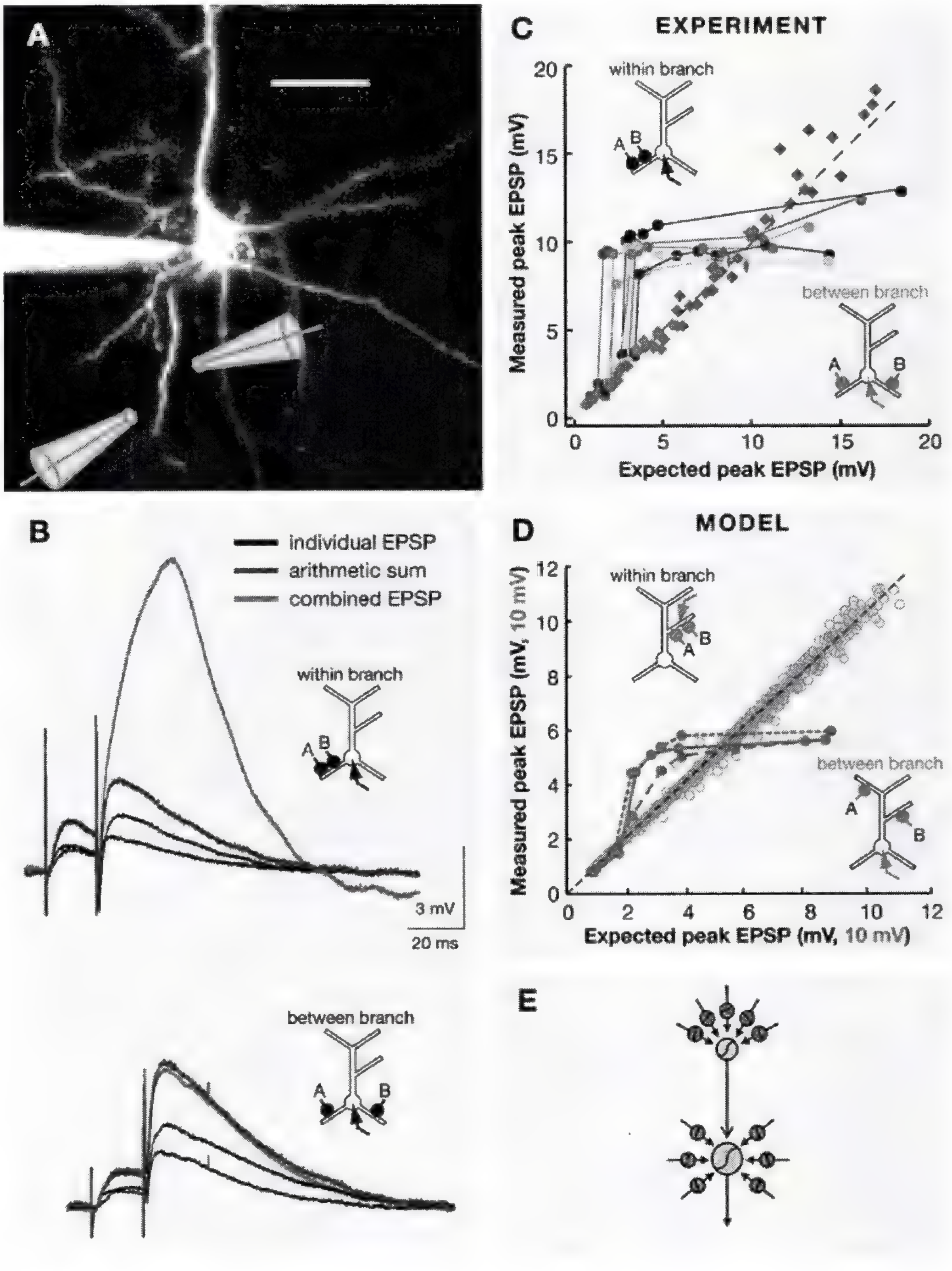
Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

coupled to, but semi-independent from, the electrical signaling in the plasma membrane (Barlow 1996). Such a "two level" arrangement could have enormous computational power, which is only beginning to be explored.

EXAMPLES OF REAL-WORLD DENDRITIC COMPUTATION

In the previous section we described how the passive and active properties of the dendrites can endow them with computational features. The key question, of course, is whether the



Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

brain takes advantage of these building blocks to perform computations. It is extremely difficult to show directly that a particular computational strategy is both necessary and sufficient to explain the computational behavior of networks. However, a few favorable cases have provided strong circumstantial evidence for dendritic computation playing a key and possibly essential role in computations performed by a neural network.

Directional Selectivity

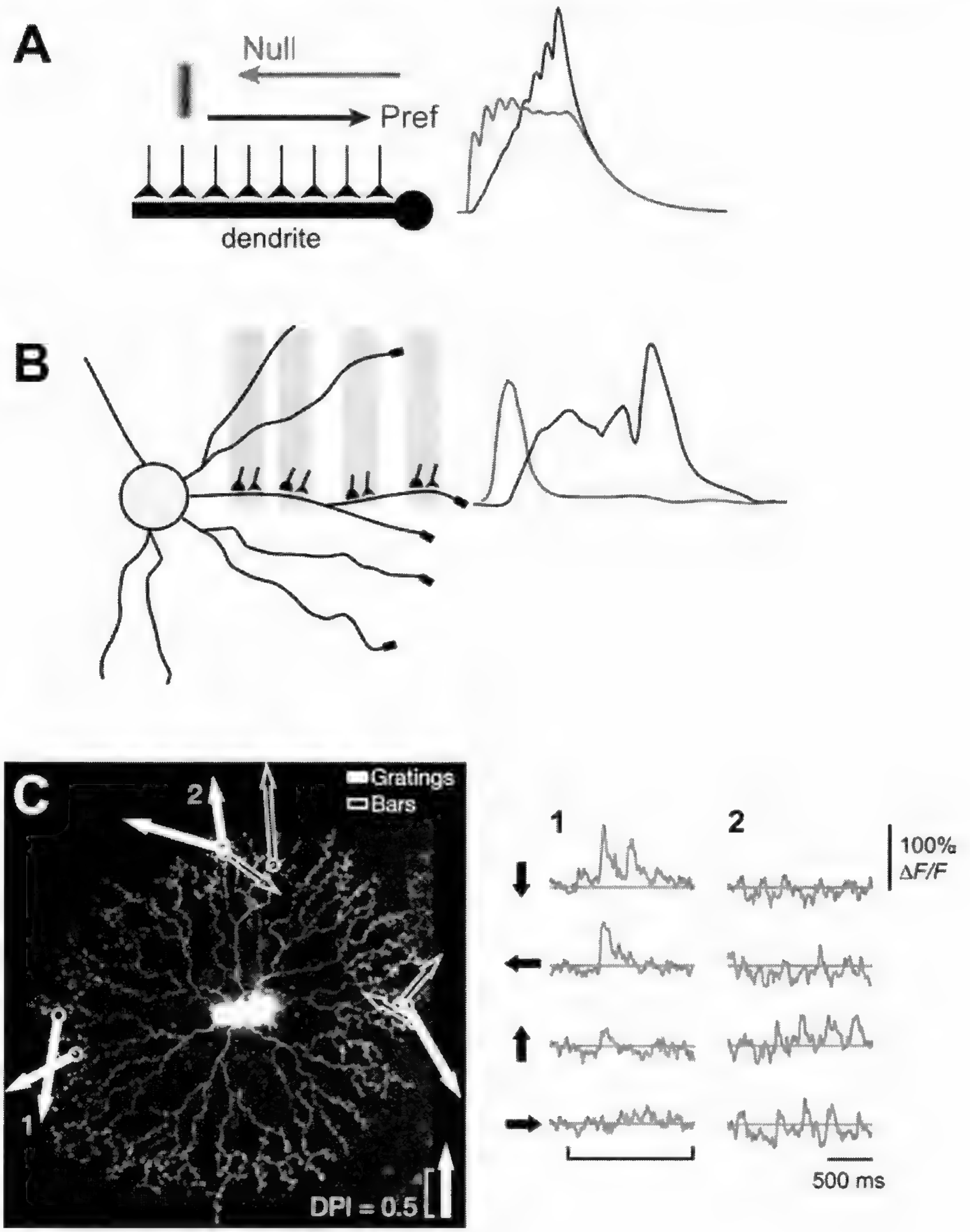
Perhaps the most extensively studied computation on the single-cell level is direction selectivity. Direction-selective neurons respond to image motion in a preferred (PREF) direction but not in the opposite NULL direction. They can be found in many species from fly eyes to mammalian cortex, and in all these cases a role for dendritic computation has been proposed. One of the first and most convincing experiments demonstrating dendritic involvement in directional selectivity was provided by Single & Borst (1998). Using imaging of dendritic calcium signals in tangential cells of the fly visual system in vivo, they showed that the input to each dendritic branch, and thus the resulting dendritic calcium signal, is already directional selective, but the dendritic filtering is required to maintain a coherent response free from spatial pattern properties of the visual scene to ensure a purely direction-selective output signal in the axon.

In this section we focus on work on the direction-selective retinal ganglion cells (DSGC) described by Barlow & Levick (Barlow et al. 1964, Barlow & Levick 1965) as a case study for dendritic computation. Rall (1964) provided the first model for how dendrites can implement a directionally selective

Figure 5

Dendritic multiplication in pyramidal cell dendritic branches. (*A*) Two stimulating electrodes were positioned near selected basal dendrites of a layer 5 pyramidal neuron filled with the calcium-sensitive dye Oregon Green BAPTA-1 (200 μM). A dendritic branch was visualized using a confocal microscope, and two stimulation electrodes were positioned in close proximity to the selected branch (*blue*). Scale bar, 75 μm. (*B*) Electrodes were activated first individually (*black traces*) and then simultaneously (*red traces*), and somatic EPSPs were recorded. Blue traces show the arithmetic sum of the two individual responses. Voltage traces are averages of four individual sweeps. Left traces show within-branch summation. The two electrodes were positioned near the same dendritic branch, separated by 20 μm (150 μm from the soma). Right traces show between-branch summation, where the two electrodes stimulated different branches and summation was linear. (*C*) Summary plot shows predicted versus actual combined responses in seven basal dendrites and one apical oblique dendrite (*pink curve*). Colored circles show sigmoidal modulation of within-branch summation (*blue and yellow*, without bicuculline to block GABAergic inhibition; *dark green trace*, with locally applied 10 μM bicuculline; *five remaining traces*, 1 μM bicuculline). Dashed line denotes exact linear summation. Green diamonds show between-branch summation experiments (12 branch pairs, 4 of them apical oblique dendrites). (*D*) Modeling data: Summation of single-pulse EPSPs in the apical oblique dendrites of a CA1 pyramidal cell model showed a similar overall pattern (Poirazi et al 2003a), including sigmoidally modulated within-branch summation (*red circles*) and linear between-branch summation (*open green circles*). Within-branch data for dendrites are attached to the apical trunk 92 μm (*short dashes*), 232 μm (*solid*), and 301 μm (*long dashes*) from the soma. Because of the uneven distances to the somatic recording electrode, recordings shown were made within the respective dendrite; for these data, axis values are scaled up 10, thus 0 mV, 20 mV, 40 mV, and so on. Modified from Polsky et al. 2004. (*E*) Schematic representation of a speculative neural network model based on the data shown in *A–C* and **Figure 4** (see also Häusser & Mel 2003). Blue branches represent the distal apical inputs, and red branches denote the basal inputs. Together, these inputs constitute the first layer of the network model, each performing supralinear summation of synaptic inputs as shown in *B* (*small circles with sigmoids*). The outputs of this first layer feed into two integration zones: one near the apical tuft (*top*) and one near the soma. These integration zones constitute the second layer of the network model (*large circles with sigmoids*).

unit (**Figure 6*A***). The idea is very simple: Synaptic input generated on the distal part of the dendrite is delayed at the soma by the dendritic filtering compared with proximal input. If synaptic inputs are activated in a sequence starting from the distal location of the dendrite toward the soma (the centripetal direction), then the EPSPs in the soma will sum effectively and the resulting somatic voltage would be large. In contrast, activating the same inputs in the centrifugal direction would result in a much lesser degree of summation because the proximal EPSP will decay by the time the distally originated EPSP will arrive. Assuming that the voltage peak is translated into action potentials, the neuron will show directional selectivity. Although the mechanism proposed is clearly feasible, there



Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

is currently little direct evidence to support this model in cases where directional selectivity has been found (Anderson et al. 1999, Euler et al. 2002).

An ongoing debate exists about identifying the earliest level of neurons that show directional selectivity. Koch et al. (1982) proposed that the nonlinear interaction between the excitation and inhibition can explain directional selectivity in retinal ganglion cells. The essential assumption was that there is an asymmetry in the spatial distribution of inhibitory and excitatory inputs to the cell such that the inhibition biased and shifted to the NULL direction. In this way when the inputs are sweeping through the receptive field in the PREF direction the excitation is acting before the inhibition and the integration of excitatory inputs cause the neuron to respond. If the input moves in the NULL direction, then the inhibition is "on the path" of the excitation and vetoes it, preventing the neuron from responding. Recently, Taylor et al. (Taylor & Vaney 2002) showed, using intracellular recording from direction-selective retinal ganglion cells (DSRGC), that indeed such asymmetry in the inhibition exists, but these results are debatable (see also Borg-Graham 2001, Fried et al. 2002, Taylor & Vaney 2000).

The other possibility is that the input to the DSRGC is already direction selective. Using similar assumptions to the model by Koch, Borg-Graham & Grzywacz (1992) (**Figure 6*B***) showed that the directional selectivity could be computed in the individual dendritic branches of starburst amacrine cells (SBAC), which are presynaptic to the retina ganglion cells. These cells do not have appropriate axons; rather each dendrite has an "output" synapse at its distal end. By using two-photon optical imaging of Ca^{2+} concentrations in the dendrites of SBAC, Euler et al. (2002) showed that indeed the Ca^{2+} concentration at the tip of the dendrites of SBAC is direction selective (**Figure 6*C***). However, in contrast to the model, the response is still selective in the presence of $GABA_A$ blockers. In summary, it seems that directional selectivity is indeed computed by individual dendrites of SBAC, but the mechanism by which it is computed is still not fully understood.

Coincidence Detection in Auditory Neurons

Another system in which the contribution of the dendrites to computation has been demonstrated is the sound localization system of chicks (Agmon-Snir et al. 1998). In this system a special type of neuron is responsible for computing the time difference between sounds arriving to the two ears. Each

Figure 6

Dendritic mechanisms for directional selectivity. (*A*) A linear model exploiting the filtering properties of passive dendritic cables. When the input advances to the right (the preferred direction), the first EPSPs are widened by dendritic filtering, which gives time for the later input to sum temporally and build a large voltage response. When the input moves in the null direction, the first large EPSP decays by the time the last EPSP arrives at the soma, and a smaller peak response is achieved. Note that this mechanism is not very robust because the difference in peak amplitude between the two scenarios is small, and the time integral of the voltage (corresponding to the total amount of charge) is identical. (*B*) A model of the starburst amacrine cell in the retina (Borg-Graham 1992). The input to each amacrine dendrite is composed of excitatory and inhibitory inputs that have symmetric receptive fields. Although the starburst cell is radially symmetric, the symmetry breaks with respect to the direction-selective circuit because the outputs to the directionally selective ganglion cells are on the distal tip of each dendrite. Furthermore, these outputs are formally direction selective, in the sense that the integral of the response depends on direction, because of the nonlinear interaction between excitation and inhibition, as described by Rall (1964) and Koch et al. (1982). (*C*) Imaging of internal Ca^{2+} concentration from dendrites of starburst amacrine cells shows that these dendrites are indeed direction selective (Euler & Denk 2002). The experiments are not consistent with the mechanism in *A*, but they do not completely agree with the model in *B* either, because blocking $GABA_A$ receptors retains some of the directional selectivity.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

neuron responds only to a very precise time difference, which corresponds to a specific location in space. The neurons contain only two major dendrites, and each dendrite receives inputs only from one ear. The inputs are supposedly arranged in such a way that there is a constant delay between the inputs arriving from one ear and the inputs arriving from the second. Coincident inputs from both ears arriving to the two dendrites are summed up at the soma and cause the neuron to emit action potentials. However, when coincident spikes arrive from the same ear, they arrive at the same dendrite and thus their summation is sublinear, resulting in a subthreshold response (**Figure 1*A***). Moreover, Rubel and colleagues (Smith & Rubel 1979) showed that there is an inverse relationship between the preferred frequency of the neurons and their dendritic length, supporting the hypothesis that the dendrites are directly contributing to the computation. This is in agreement with the model because for high-frequency inputs the dendritic filtering causes accidental spikes that are out of phase to summate and cross the threshold. Thus the advantage of the dendrites in the low-frequency range becomes a burden in the high-frequency range, and the auditory coincidence detection neuron is better off with shorter dendrites.

Temporal Integration Over Long Timescales

All the computations discussed above take place on relatively short timescales; the neuron responds almost instantaneously to the input. Computations over longer timescales, such as those required for working memory, are usually attributed to network phenomena or involving molecular dynamics. In this context, computation of time integration becomes a challenging problem. How does a system integrate transients and maintain the computed integral for a long period, far longer than its intrinsic time constant? One example of such a system is the oculomotor system in the goldfish, where neurons maintain stable firing rates corresponding to the position of the eye and switch between them rapidly when saccades occur. Although previous work has focused largely on network explanations for such forms of time integration, dendritic mechanisms (Brody et al. 2003) in single neurons may also contribute. In particular, one intrinsic mechanism that could be involved in such a phenomenon is bistability. This is a property of a dynamical system exhibiting more than one stable point. Once driven to approach a stable point, it stays there. Neurons expressing specific types of voltage-gated ionic currents, for example, can show bistability. Some indirect evidence for dendritic bistability exists (Booth & Rinzel 1995, Milojkovic et al. 2004). Recently, Loewenstein & Sompolinsky (2003) presented a model in which a specifically dendritic bistability enables the dendrite to become a time integrator. In this model, the dendritic concentration of Ca^{2+} is bistable. If at one end of the dendrite it is forced to be in the up state, and at the other edge to be in the down state, then a standing wave of Ca^{2+} concentration is created. The location of the wavefront represents the result of the integral such that incoming transient synaptic input repositions the location of the front. As long as no input arrives, the front will keep its position and the dendrite will preserve the "working memory." A related study, based on an earlier model (Rosen 1972), has been proposed on the basis of dendritic bistability involving voltage-sensitive conductances (Goldman et al. 2003).

Image Processing in Dendrites of Fly Neurons

A key component of visual information processing by the fly nervous system is convolution, whereby an image is smoothed to remove noise and improve extraction of salient features. This operation is thought to be performed by a group of neurons in the lobula plate in the third visual ganglion known as the horizontally sensitive tangential cells. These

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

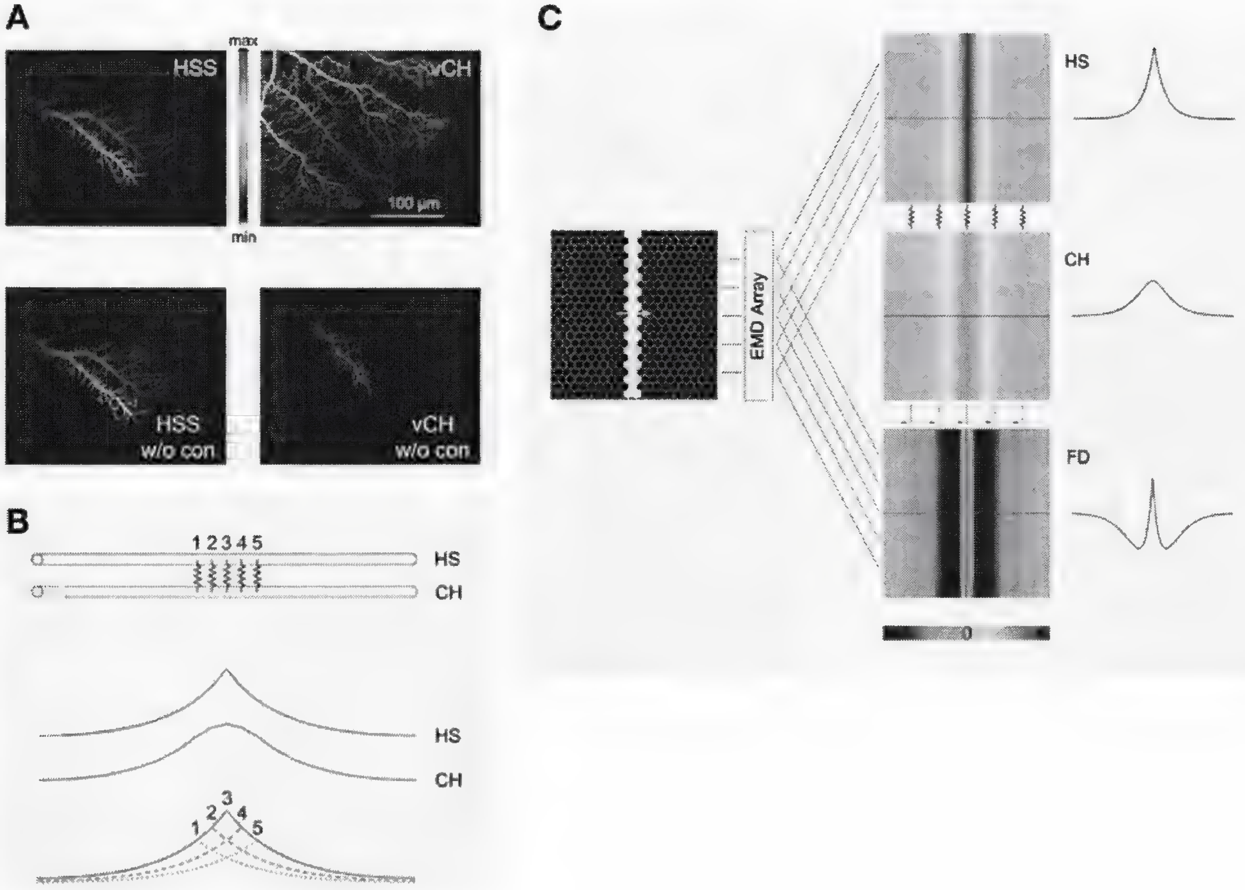


Figure 7

Image processing in visual interneurons of the fly. (*A*) *Top*: Spread of membrane potential in a model of neurons, which are horizontally sensitive of the southern area of the visual field (HSS) (*left*) and ventral centrifugal horizontal cells (vCH) model (*right*) after local current injection into HSS. *Bottom*: same as top panels, but HSS and vCH models were not connected to each other. Current was injected in HSS (*left*) and vCH (*right*). (*B*) Simplified model of HS and CH neurons. *Top*: Two cylinders (HS and CH) are connected by five linear conductances surrounding the location of current injection. *Middle panel*: in HS, the signal spreads with an exponential decay. The CH spread is broader. *Bottom panel*: The CH spread can be approximated by the sum of passive spread through each conductance. (*C*) Consequences of the CH cell dendritic image blurring for relative motion detection. An array of elementary motion detectors computes the image motion in a retinotopic way, feeding onto the dendrites of HS and figure-detection (FD) cells. This motion representation is blurred in the dendrites of the CH cell via dendro-dendritic connections between HS and CH cells. By conveying inhibitory dendro-dendritic input to FD cells, being subtracted from the retinotopic input, an enhancement of the motion edges is achieved. Modified from Cuntz et al. 2003.

neurons respond to visual motion in a directionally selective way and can be divided into two groups of neurons: the horizonal system (HS) cells and the centrifugal horizonal (CH) cells. Specific CH cells, the ventral CH (vCH) cells, are electrically coupled via dendritic gap junctions to HS cells. Retinotopic input to HS cells is already filtered by the electrotonic decay of membrane potential in HS cell dendrites. As shown in **Figure 7** (Cuntz et al. 2003), the coupling via gap junctions then imposes the filtered membrane potential of the HS cell onto the CH dendritic tree, where another round of low-pass filtering takes place. The end result is a blurred and de-noised version of the original image.

Thus, the biophysical properties of dendritic trees can implement non-trivial image processing operations in a simple and elegant manner.

Looming Sensitive Neurons in the Locust

The lobula giant movement detector (LGMD) is an identified neuron in the locust visual system whose output firing rate is most sensitive to objects approaching a collision course (looming visual stimuli), indicating a forthcoming collision (**Figure 8**). The timing of the peak firing rate comes with a fixed delay after the time at which a looming object reaches a fixed-threshold angular size, on the retina, independent of the object's approach speed or size (Gabbiani et al. 1999, 2004). A mathematical model supported by experimental results predicts that this behavior could be explained as a multiplication of two parameters of the approaching object, namely its angular size and speed of approach (Gabbiani et al. 2002). The LGMD neuron has a unique dendritic structure composed of a fan-like tree and two additional sub trees. The synaptic input is segregated such that excitatory, motion-sensitive input arrives on the major fan-like tree, and inhibitory size-sensitive inputs arrive in separated ON/OFF channels at the remaining two subtrees. The multiplication is thought to be implemented such that each of the relevant parameters is encoded logarithmically in one of these subtrees, and the dendritic sum of the two types of inputs results with the sum of logarithm (corresponding to the logarithm of the multiplication). The spiking mechanism on this combined input implements an approximate exponentiation that inverses the logarithm, and the result of the multiplication is thus encoded in the firing rate. The accessibility of the LGMD neuron to dendritic recordings and optical imaging in vivo makes it a promising candidate for understanding the biophysics of a high-level computation in dendrites in the near future.

Forward Masking of Cricket Songs

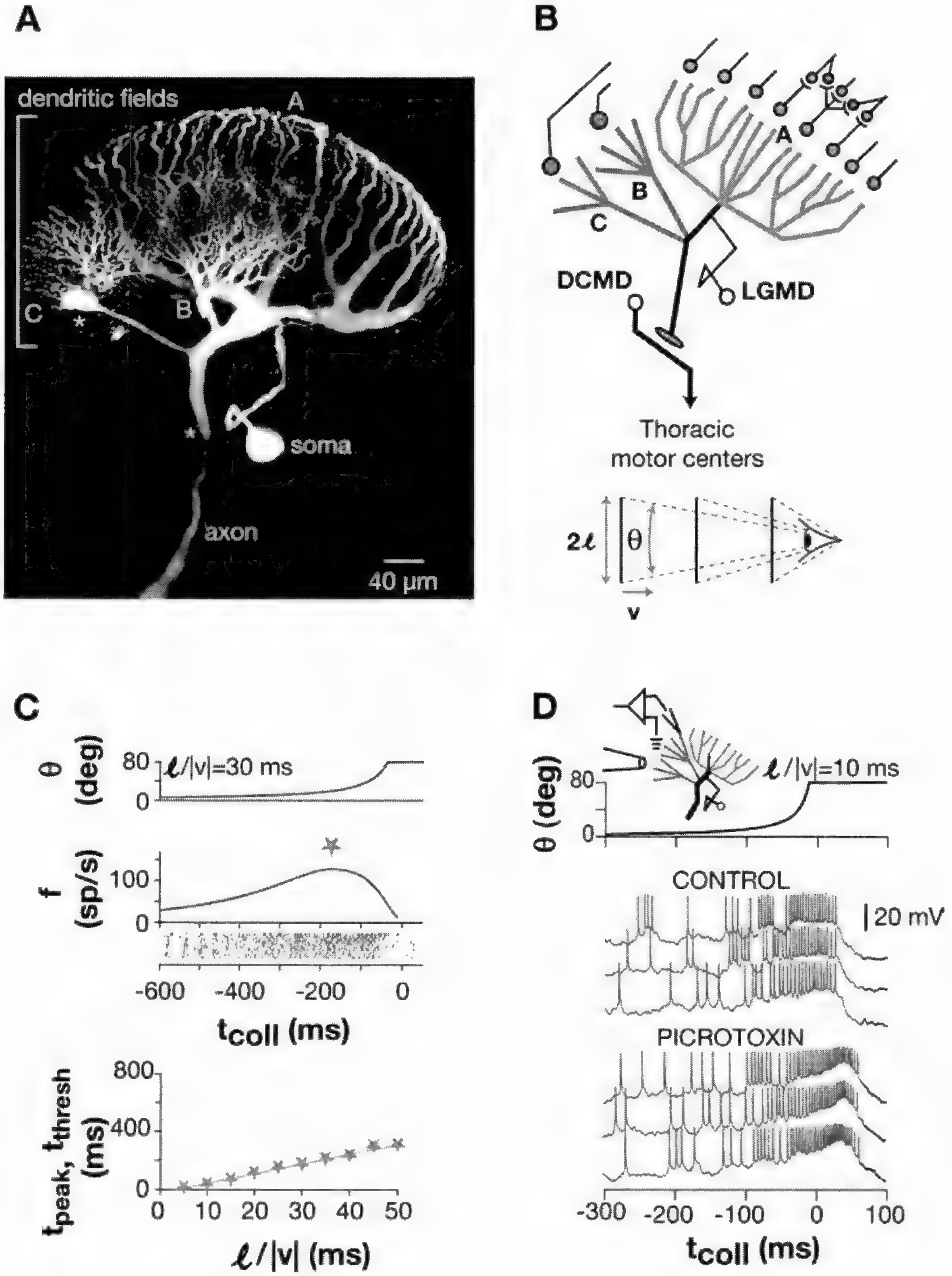
Omega neurons in female crickets respond to the male calling song with bursts of action potentials. The response of an omega to a particular sound can be dramatically attenuated if it is preceded by an identical but louder sound. This suppression is known as

Figure 8

Computation in dendrites of locust looming-sensitive neurons. (*A*) The LGMD neuron's dendritic tree consists of three distinct subfields (*A–C*). Subfield *A* receives motion-sensitive excitatory inputs, whereas subfields *B* and *C* receive phasic inhibition related to object size. As in many invertebrate neurons the soma lies outside the electrical signal propagation path, and spikes are generated at the point where the axon is thinnest. (*B*) Schematic illustration of the neuronal inputs received by the LGMD. Postsynaptic inhibitory regions of the LGMD are illustrated in red and excitatory ones in green. Green and red dots represent inhibitory and excitatory synapses, respectively. Yellow dots indicate cell bodies. *Bottom*: A solid object of size 2ℓ is approaching the animal on a direct collision course with a constant velocity v. The angle subtended by the object is represented by θ. (*C*) Top panel shows the time course of θ for the looming stimulus, and the middle panel shows individual spike trains of the LGMD neuron in response to repetitions of this stimulus. The blue line above the spike trains represents the average instantaneous firing rate, and its peak is marked by a star. *Bottom*: The relation between the peak firing time relative to collision as a function of $\ell/|v|$ is nearly linear. This is equivalent to the angular size subtended by the object being a fixed constant δ ms prior to the peak, independent of the stimulation. This angular size is typically in the range of 15°–35°. Thus, LGMD's peak firing time acts as an angular threshold detector. (*D*) Top three traces are intracellular dendritic recordings in response to a looming stimulus, and bottom three traces show the response after picrotoxin injection to the lobula. Picrotoxin prolongs the responses, and the peak firing rate no longer predicts collisions. Adapted from Gabbiani et al. 2004.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

forward masking and represents a simple form of gain control, allowing the female cricket to focus on the loudest male in the presence of multiple competing males. Sobel & Tank (1994) have shown that biochemical dendritic signaling underlies this simple computation. By imaging dendritic calcium signals in omega neurons, they showed that a loud simulated calling song triggers a large, long-lasting dendritic calcium transient associated with a potassium conductance that suppresses excitability. The time course of the calcium signal was tightly correlated with the time course of forward masking. To demonstrate



Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

a causal relationship, Sobel & Tank showed that buffering dendritic calcium prevents the hyperpolarization and reduction in excitability. They also demonstrated that uncaging calcium in the dendrite produces a similar hyperpolarization and dampening of excitability to that generated by the sound. These findings demonstrate how dendritic mechanisms—conversion of action potentials into a calcium signal and then activation of a potassium conductance—can be directly linked to a computational task relevant to behavior.

Dendritic Mechanisms and Behavior

Two recent pioneering experimental studies have opened the door for exploring dendritic function of mammalian neurons in an entirely new framework. By using molecular techniques to manipulate dendritic ion channels in transgenic animals, Nolan et al. (2004) and Bernard et al. (2004) have demonstrated that it is possible to link the excitable properties of distal dendrites with network-level and behavioral phenomena. Bernard et al. (2004) demonstrated that in an animal model of temporal lobe epilepsy, the excitability of CA1 pyramidal cell dendrites is enhanced by downregulation of dendritic A-type K^+ channels by phosphorylation, combined with reduced expression of these channels. Although the evidence remains indirect, the enhanced dendritic excitability associated with this channelopathy may contribute to the observed reduction in seizure threshold in the hippocampus.

Nolan and colleagues (2004) generated a transgenic mouse with a forebrain-restricted deletion of the HCN1 gene, which encodes the hyperpolarization-activated cation current I_h. These mice exhibit enhanced performance in hippocampal-dependent learning and memory tasks. On the cellular and network level, the mice demonstrate enhanced theta oscillations and LTP. Because previous work has demonstrated that HCN1 channels are highly concentrated in the distal dendrites of CA1 pyramidal neurons (Magee 1998), the authors examined integration of proximal and distal synaptic input to CA1 pyramidal neurons. They demonstrated that distal perforant path synaptic inputs are selectively enhanced in the HCN1 knockout and that LTP of these inputs is also enhanced, whereas LTP at the more proximal Schaffer collateral input is unchanged. These results provide some of the best available evidence linking integration of distal dendritic synaptic input with behavior and point to the importance of independent regulation of excitability in subcompartments of the dendritic tree.

CHALLENGES FOR THE FUTURE: A WISH LIST FOR DENDRITIC COMPUTATION

The ultimate challenge for those interested in dendritic computation is to show that computation conveys a significant advantage in the operation of real neural circuits. This advantage is difficult to show directly, given that dendrites also have other properties not directly related to computation, such as their important structural role in determining brain wiring (Chklovskii 2004). Here we outline several of the key challenges faced by experimenters working at different levels to understand the contribution of dendrites to computation in the mammalian brain.

For Molecular Biologists: Designer Dendrites

To manipulate the computational properties of dendritic trees, we want to be able to manipulate dendritic shape and the spatial distribution and properties of voltage-gated channels using genetic tools. The molecular regulation of dendritic growth and shape has become a burgeoning field over the past decade, and many kinases have been identified that can be regulated to produce changes in dendritic form (Scott & Luo 2001). In parallel, our understanding of ion channel trafficking

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

regulation and the local synthesis of ion channels in dendrites has also made substantial progress (Horton & Ehlers 2003, Misonou et al. 2004, Misonou & Trimmer 2004). Within the next decade it should be possible, using transgenic and viral techniques, to selectively modify single neurons or entire populations of neurons to generate dendritic subunits with defined shapes and electrical properties. Such "designer dendrites" can be used to identify the role of specific dendritic mechanisms, such as action potential backpropagation, for the function and formation of specific neural networks and ultimately can serve as a well-defined bridge between molecules and behavior.

Many dendritic mechanisms require that specific kinds of synaptic input are addressed to specific regions of the dendritic tree or furthermore require that synaptic inputs are highly spatially clustered. Although there is already considerable evidence from decades of anatomical work that precise spatial targeting of certain kinds of synaptic inputs is achieved in some cell types (Freund & Buzsaki 1996, Somogyi et al. 1998), identifying the molecular mechanisms responsible for such targeting will be very important for understanding how synaptic connectivity interacts with and defines dendritic computational mechanisms (Ango et al. 2004). It will be of great interest to harness these mechanisms to identify and manipulate the spatial arrangement of specific types of synaptic inputs to the dendritic tree. These mechanisms could be exploited first to label inputs selectively conveying different streams of information. This will allow us to test directly to what extent inputs carrying similar or divergent information are clustered on neighboring stretches of dendrite, or whether targeting is essentially random on the scale of microns to dozens of microns. Second, once patterns of spatial clustering have been identified, then by manipulating molecular mechanisms it may be possible to disrupt or redirect such clustering in a defined way to permit causal links to be made between the

HOW CAN WE PROVE THAT DENDRITES ARE INVOLVED IN COMPUTATION?

Proving that dendrites are both necessary and sufficient for a particular computation relevant to behavior is a very difficult challenge. Necessity is within reach for the relatively simple case where the computation is accomplished by a single identified neuron, as is the case for some invertebrate sensory neurons (see text) or when the underlying biophysical mechanism depends on a single channel type. Demonstrating sufficiency is much more difficult, particularly because dendritic computation is a process that is tightly interlinked with the proper functioning of the entire system. Nevertheless, we outline here a list of objectives that must be achieved to prove that dendrites are required for computation. These should not necessarily be addressed in a linear sequence; rather, it will be beneficial to attack these problems in parallel.

Identify the Computation:

Probing the contribution of dendrites to computation is possible only when the computation of the neuron bearing the dendrites is identified. This requires identifying a simple behavior that involves a recognizable kind of computation (e.g., filtering, convolution, pattern recognition) and tracing it to the neurons responsible.

Defining the Mechanism:

Use recordings (e.g., electrophysiological or imaging) from dendrites of these neurons in an accessible preparation (e.g., brain slices) to define the dendritic signals and biophysical mechanisms that may underlie the behavior.

Correlation in the Intact Preparation:

Use recordings from dendrites in an intact preparation to show strong correlations between dendritic signals linked with the identified computation and the behavior of the animal.

Manipulation to Define a Causal Link:

Manipulate a dendritic mechanism to determine if it is both necessary and sufficient to explain the computation. Selectively knock out the mechanism and demonstrate that the behavior is impaired. Activate or modify the dendritic mechanism to demonstrate that the behavior is modified in the expected direction.

Modeling the Computation:

Use modeling to define an algorithm that describes the computation, or sequence of computations, performed by the dendrites that can plausibly explain the behavior. Modeling of single neurons and neural networks can also be used to confirm that the computation can convey a significant benefit (which can help to establish sufficiency).

local nonlinear computations described above and behavior.

bAP: backpropagating action potential

CH: centrifugal horizontal

DSGC: direction-selective retinal ganglion cells

EPSP: excitatory postsynaptic potential

GABA_A: γ-aminobutyric acid type A

HS: horizontal system

LGMD: lobula giant movement detector

NMDA: N-methyl-D-aspartate

SBAC: starburst amacrine cells

VCH: ventral centrifugal horizontal

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

For Neurophysiologists: Putting Dendrites Back into the Network

Most of our understanding of dendritic function has come from studies in isolated preparations (brain slices and culture preparations). Although this approach has been very successful in defining the biophysical basis of dendritic excitability and identifying computational subunits in neurons, it is associated with two main problems. First, the baseline conditions in these preparations are often very different from those pertaining in the intact brain, where the presence of high levels of background synaptic input can fundamentally change dendritic processing (Destexhe et al. 2003, Williams 2004). Second, the technical limitations imposed by slice experiments, together with uncertainty about the spatiotemporal pattern of synaptic inputs to single neurons in the intact brain, have made it difficult to identify which of the many mechanisms studied in vitro are also operating in vivo and may actually be relevant for computation in the intact brain. Fortunately, it is now possible to investigate dendritic function directly in vivo. Both electrophysiological (Buzsaki & Kandel 1998, Kamondi et al. 1998, Larkum & Zhu 2002, Loewenstein et al. 2005) and optical (Helmchen et al. 1999, Svoboda et al. 1999, Waters et al. 2004) techniques now exist for recording electrical and chemical dendritic signaling in anesthetized and awake, head-restrained animals. The prospect of using two-photon imaging techniques to monitor dendritic signaling in freely moving animals (Helmchen et al. 2001) should permit correlations to be established between dendritic events associated with computation (e.g., dendritic spikes) and behavior. In combination with the molecular tools outlined above and new tools for noninvasively manipulating neuronal activity in defined populations (Fetcho & Bhatt 2004, Lima & Miesenböck 2005), it should be possible to identify the causal links between dendritic computation and behavior.

For the Theorist: Proving the Benefits of Dendritic Computation

Ultimately, understanding the role of dendrites in neural computation requires a theory. This theory must identify the benefits of having dendrites and reveal the basic principles used to provide these benefits. To make advances toward such a theory, efforts should be made in three major directions. First, we need to construct algorithms based on the dendritic toolkit and show how specific computations can be achieved, making predictions that are experimentally testable (e.g., Agmon-Snir et al. 1998). Second, given that realistic modeling of single neurons has reached a relatively mature phase, we need to take advantage of the capability of such models to simulate conditions that are very difficult to test experimentally. One important task is to explore how different components of the dendritic toolkit interact with each other (e.g., how do local interactions between excitation and inhibition influence local dendritic spikes?). It is also essential to use these models to see how realistic conditions, such as noise, neuromodulation, and adaptation, affect the computational properties of the dendrites. It is a major challenge to understand how stability of the algorithmic computation can be maintained in the face of these variables, such that it is resistant to them or such that these variables can be synergistically exploited. The third challenge is to put dendrites back into networks. This will be greatly aided by the construction of reduced models of dendritic neurons (e.g., Rall 1964, Pinsky & Rinzel 1994) which capture essential features of dendritic function that could be exploited for computation. The ultimate step will be to build artificial neural networks incorporating such reduced models of the single neuron to demonstrate to what degree dendritic algorithms enhance the performance of neural networks in well-defined tasks.

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

CONCLUSIONS

Although dendrites have been studied for decades, the field of dendritic computation is still in its infancy. This is partly because dendrites remain relatively inaccessible and have only recently begun to yield their secrets to the onslaught of multiple new experimental tools. However, the real challenge is a deeper one, faced by many areas of neuroscience (and biology in general): how to evaluate the importance of mechanisms on the molecular and cellular level for computation at the behavioral level. The ability not only to record electrical and chemical signals in the intact brain but also to manipulate the structure and function of dendrites using molecular tools will hopefully allow us to move from the descriptive level, correlating dendritic signals linked to computation with behavior, toward directly testing the causal nature of these links. Such experiments will provide a deeper understanding of how single neurons contribute to computation in the brain and should inspire the development of novel neural network architectures with the computational powers of real brains.

ACKNOWLEDGMENTS

We are grateful to Peter Dayan, Lyle Graham, Julian Jack, Bartlett Mel, Arnd Roth, and Idan Segev for many helpful discussions and for comments on the manuscript. We thank the HFSP, Gatsby Foundation, and Wellcome Trust for financial support.

LITERATURE CITED

Agmon-Snir H, Carr CE, Rinzel J. 1998. The role of dendrites in auditory coincidence detection. *Nature* 393:268–72

Anderson JC, Binzegger T, Kahana O, Martin KA, Segev I. 1999. Dendritic asymmetry cannot account for directional responses of neurons in visual cortex. *Nat. Neurosci.* 2:820–24

Ango F, di Cristo G, Higashiyama H, Bennett V, Wu P, Huang ZJ. 2004. Ankyrin-based subcellular gradient of neurofascin, an immunoglobulin family protein, directs GABAergic innervation at purkinje axon initial segment. *Cell* 119:257–72

Ariav G, Polsky A, Schiller J. 2003. Submillisecond precision of the input-output transformation function mediated by fast sodium dendritic spikes in basal dendrites of CA1 pyramidal neurons. *J. Neurosci.* 23:7750–58

Barlow HB. 1996. Intraneuronal information processing, directional selectivity and memory for spatio-temporal sequences. *Network* 7:251–59

Barlow HB, Hill RM, Levick WR. 1964. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *J. Physiol.* 173:377–407

Barlow HB, Levick WR. 1965. The mechanism of directionally selective units in rabbit's retina. *J. Physiol.* 178:477–504

Bernard C, Anderson A, Becker A, Poolos NP, Beck H, Johnston D. 2004. Acquired dendritic channelopathy in temporal lobe epilepsy. *Science* 305:532–35

Blomfield S. 1974. Arithmetical operations performed by nerve cells. *Brain Res.* 69:115–24

Booth V, Rinzel J. 1995. A minimal, compartmental model for a dendritic origin of bistability of motoneuron firing patterns. *J. Comput. Neurosci.* 2:299–312

Borg-Graham LJ. 2001. The computation of directional selectivity in the retina occurs presynaptic to the ganglion cell. *Nat. Neurosci.* 4:176–83

Borg-Graham LJ, Grzywacz NM. 1992. A model of the directional selectivity circuit in retina: transformation by neuron singly and in concert. In *Single Neuron Computation*, ed. T McKenna, J Davis, SF Zornetzer, pp. 347–76. San Diego: Academic

Brody CD, Romo R, Kepecs A. 2003. Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations. *Curr. Opin. Neurobiol.* 13:204–11

Buzsaki G, Kandel A. 1998. Somadendritic backpropagation of action potentials in cortical pyramidal cells of the awake rat. *J. Neurophysiol.* 79:1587–91

Cai X, Liang CW, Muralidharan S, Kao JP, Tang CM, Thompson SM. 2004. Unique roles of SK and Kv4.2 potassium channels in dendritic integration. *Neuron* 44:351–64

Cajal SR. 1911. *Histologie du Système Nerveux de l'Homme et des Vertébrés*. Paris: Maloine

Carruth M, Magee JC. 1999. Dendritic voltage-gated ion channels regulate the action potential firing mode of hippocampal CA1 pyramidal neurons. *J. Neurophysiol.* 82:1895–901

Cauller LJ, Connors BW. 1992. Functions of very distal dendrites: experimental and computational studies of layer I synapses on neocortical pyramidal cells. In *Single Neuron Computation*, ed. T McKenna, J Davis, SF Zornetzer, pp. 199–229. Boston: Academic

Chklovskii DB. 2004. Synaptic connectivity and neuronal morphology: two sides of the same coin. *Neuron* 43:609–17

Chklovskii DB, Mel BW, Svoboda K. 2004. Cortical rewiring and information storage. *Nature* 431:782–88

Cook EP, Johnston D. 1997. Active dendrites reduce location-dependent variability of synaptic input trains. *J. Neurophysiol.* 78:2116–28

Cook EP, Johnston D. 1999. Voltage-dependent properties of dendrites that eliminate location-dependent variability of synaptic input. *J. Neurophysiol.* 81:535–43

Cuntz H, Haag J, Borst A. 2003. Neural image processing by dendritic networks. *Proc. Natl. Acad. Sci. USA* 100:11082–85

Destexhe A, Rudolph M, Pare D. 2003. The high-conductance state of neocortical neurons in vivo. *Nat. Rev. Neurosci.* 4:739–51

Doiron B, Laing C, Longtin A, Maler L. 2002. Ghostbursting: a novel neuronal burst mechanism. *J. Comput. Neurosci.* 12:5–25

Emptage N, Bliss TV, Fine A. 1999. Single synaptic events evoke NMDA receptor-mediated release of calcium from internal stores in hippocampal dendritic spines. *Neuron* 22:115–24

Euler T, Denk W. 2001. Dendritic processing. *Curr. Opin. Neurobiol.* 11:415–22

Euler T, Detwiler PB, Denk W. 2002. Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature* 418:845–52

Fatt P, Katz B. 1953. The effect of inhibitory nerve impulses on a crustacean muscle fibre. *J. Physiol.* 121:374–89

Fetcho JR, Bhatt DH. 2004. Genes and photons: new avenues into the neuronal basis of behavior. *Curr. Opin. Neurobiol.* 14:707–14

Fetz EE, Gustafsson B. 1983. Relation between shapes of post-synaptic potentials and changes in firing probability of cat motoneurones. *J. Physiol.* 341:387–410

Freund TF, Buzsáki G. 1998. Interneurons of the hippocampus. *Hippocampus* 6:347–470

Fried SI, Münch TA, Werblin FS. 2002. Mechanisms and circuitry underlying directional selectivity in the retina. *Nature* 420:411–14

Gabbiani F, Krapp HG, Hatsopoulos N, Mo CH, Koch C, Laurent G. 2004. Multiplication and stimulus invariance in a looming-sensitive neuron. *J. Physiol. Paris* 98:19–34

Gabbiani F, Krapp HG, Koch C, Laurent G. 2002. Multiplicative computation in a visual neuron sensitive to looming. *Nature* 420:320–24

Gabbiani F, Krapp HG, Laurent G. 1999. Computation of object approach by a wide-field, motion-sensitive neuron. *J. Neurosci.* 19:1122–41

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Gasparini S, Migliore M, Magee JC. 2004. On the initiation and propagation of dendritic spikes in CA1 pyramidal neurons. *J. Neurosci.* 24:11046–56

Golding NL, Spruston N. 1998. Dendritic sodium spikes are variable triggers of axonal action potentials in hippocampal CA1 pyramidal neurons. *Neuron* 21:1189–200

Golding NL, Staff NP, Spruston N. 2002. Dendritic spikes as a mechanism for cooperative long-term potentiation. *Nature* 418:326–31

Goldman MS, Levine JH, Major G, Tank DW, Seung HS. 2003. Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cereb. Cortex* 13:1185–95

Grutzendler J, Kasthuri N, Gan WB. 2002. Long-term dendritic spine stability in the adult cortex. *Nature* 420:812–16

Häusser M, Mel B. 2003. Dendrites: bug or feature? *Curr. Opin. Neurobiol.* 13:372–83

Häusser M, Spruston N, Stuart GJ. 2000. Diversity and dynamics of dendritic signaling. *Science* 290:739–44

Hebb D. 1949. *The Organization of Behavior*. New York: Wiley

Helmchen F, Fee MS, Tank DW, Denk W. 2001. A miniature head-mounted two-photon microscope. High-resolution brain imaging in freely moving animals. *Neuron* 31:903–12

Helmchen F, Imoto K, Sakmann B. 1996. Ca^{2+} buffering and action potential-evoked Ca^{2+} signaling in dendrites of pyramidal neurons. *Biophys. J.* 70:1069–81

Helmchen F, Svoboda K, Denk W, Tank DW. 1999. In vivo dendritic calcium dynamics in deep-layer cortical pyramidal neurons. *Nat. Neurosci.* 2:989–96

Hoffman DA, Magee JC, Colbert CM, Johnston D. 1997. K^+ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons. *Nature* 387:869–75

Holtmaat AJ, Trachtenberg JT, Wilbrecht L, Shepherd GM, Zhang X, et al. 2005. Transient and persistent dendritic spines in the neocortex in vivo. *Neuron* 45:279–91

Horton AC, Ehlers MD. 2003. Neuronal polarity and trafficking. *Neuron* 40:277–95

Iansek R, Redman SJ. 1973. The amplitude, time course and charge of unitary excitatory post-synaptic potentials evoked in spinal motoneurone dendrites. *J. Physiol.* 234:665–88

Jack JJB, Noble D, Tsien RY. 1975. *Electric Current Flow in Excitable Cells*. Oxford, UK: Oxford Univ. Press

Johnston D, Hoffman DA, Magee JC, Poolos NP, Watanabe S, et al. 2000. Dendritic potassium channels in hippocampal pyramidal neurons. *J. Physiol.* 525:75–81

Kamondi A, Acsady L, Buzsaki G. 1998. Dendritic spikes are enhanced by cooperative network activity in the intact hippocampus. *J. Neurosci.* 18:3919–28

Kepecs A, Wang XJ, Lisman J. 2002. Bursting neurons signal input slope. *J. Neurosci.* 22:9053–62

Koch C, Poggio T, Torre V. 1982. Retinal ganglion cells: a functional interpretation of dendritic morphology. *Phil. Trans. R. Soc. London* 298:227–64

Koch C, Poggio T, Torre V. 1983. Nonlinear interactions in a dendritic tree: localization, timing, and role in information processing. *Proc. Natl. Acad. Sci. USA* 80:2799–802

Koch C, Segev I. 2000. The role of single neurons in information processing. *Nat. Neurosci.* 3(Suppl.):1171–77

Kreiman G, Koch C, Fried I. 2000. Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat. Neurosci.* 3(9):946–53

Larkum ME, Kaiser KM, Sakmann B. 1999a. Calcium electrogenesis in distal apical dendrites of layer 5 pyramidal cells at a critical frequency of back-propagating action potentials. *Proc. Natl. Acad. Sci. USA* 96:14600–4

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Larkum ME, Watanabe S, Nakamura T, Lasser-Ross N, Ross WN. 2003. Synaptically activated Ca^{2+} waves in layer 2/3 and layer 5 rat neocortical pyramidal neurons. *J. Physiol.* 549:471–88

Larkum ME, Zhu JJ. 2002. Signaling of layer 1 and whisker-evoked Ca^{2+} and Na^{+} action potentials in distal and terminal dendrites of rat neocortical pyramidal neurons in vitro and in vivo. *J. Neurosci.* 22:6991–7005

Larkum ME, Zhu JJ, Sakmann B. 1999b. A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* 398:338–41

Larkum ME, Zhu JJ, Sakmann B. 2001. Dendritic mechanisms underlying the coupling of the dendritic with the axonal action potential initiation zone of adult rat layer 5 pyramidal neurons. *J. Physiol.* 533:447–66

Lima SQ, Miesenböck G. 2005. Remote control of fly behavior through genetically targeted photostimulation of neurons. *Cell.* In press

Linden D. 1999. The return of the spike: postsynaptic action potentials and the induction of LTP and LTD. *Neuron* 22:661–66

Liu G. 2004. Local structural balance and functional interaction of excitatory and inhibitory synapses in hippocampal dendrites. *Nat. Neurosci.* 7:373–79

Loewenstein Y, Sompolinsky H. 2003. Temporal integration by calcium dynamics in a model neuron. *Nat. Neurosci.* 6:961–67

Loewenstein Y, Mahon S, Chadderton P, Kitamura K, Sompolinsky H, et al. 2005. Bistability of cerebellar Purkinje cells modulated by sensory stimulation. *Nat. Neurosci.* 8:202–11

Lörincz A, Notomi T, Tamás G, Shigemoto R, Nusser Z. 2002. Polarized and compartment-dependent distribution of HCN1 in pyramidal cell dendrites. *Nat. Neurosci.* 5:1185–93

Maeda H, Ellis-Davies GC, Ito K, Miyashita Y, Kasai H. 1999. Supralinear Ca^{2+} signaling by cooperative and mobile Ca^{2+} buffering in Purkinje neurons. *Neuron* 24:989–1002

Magee J, Hoffman D, Colbert C, Johnston D. 1998. Electrical and calcium signaling in dendrites of hippocampal pyramidal neurons. *Annu. Rev. Physiol.* 60:327–46

Magee JC. 1998. Dendritic hyperpolarization-activated currents modify the integrative properties of hippocampal CA1 pyramidal neurons. *J. Neurosci.* 18:7613–24

Magee JC. 2000. Dendritic integration of excitatory synaptic inputs. *Nat. Rev. Neurosci.* 1:181–90

Magee JC, Johnston D. 1997. A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. *Science* 275:209–13

Mainen ZF. 1999. Functional plasticity at dendritic synapses. In *Dendrites,* ed. G Stuart, N Spruston, M Häusser, pp. 310–38. Oxford, UK: Oxford Univ. Press

Mainen ZF, Sejnowski TJ. 1996. Influence of dendritic structure on firing pattern in model neocortical neurons. *Nature* 382:363–66

McCulloch WS, Pitts WH. 1943. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* 5:115–33

Mehta MR. 2004. Cooperative LTP can map memory sequences on dendritic branches. *Trends Neurosci.* 27:69–72

Mel BW. 1993. Synaptic integration in an excitable dendritic tree. *J. Neurophysiol.* 70:1086–101

Migliore M, Shepherd GM. 2002. Emerging rules for the distributions of active dendritic conductances. *Nat. Rev. Neurosci.* 3:362–70

Milojkovic BA, Radojicic MS, Goldman-Rakic PS, Antic SD. 2004. Burst generation in rat pyramidal neurones by regenerative potentials elicited in a restricted part of the basilar dendritic tree. *J. Physiol.* 558:193–211

Misonou H, Mohapatra DP, Park EW, Leung V, Zhen D, et al. 2004. Regulation of ion channel localization and phosphorylation by neuronal activity. *Nat. Neurosci.* 7:711–18

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Misonou H, Trimmer JS. 2004. Determinants of voltage-gated potassium channel surface expression and localization in Mammalian neurons. *Crit. Rev. Biochem. Mol. Biol.* 39:125–45

Mizrahi A, Katz LC. 2003. Dendritic stability in the adult olfactory bulb. *Nat. Neurosci.* 6:1201–7

Nakamura T, Barbara JG, Nakamura K, Ross WN. 1999. Synergistic release of Ca^{2+} from IP3-sensitive stores evoked by synaptic activation of mGluRs paired with backpropagating action potentials. *Neuron* 24:727–37

Nakamura T, Lasser-Ross N, Nakamura K, Ross WN. 2002. Spatial segregation and interaction of calcium signalling mechanisms in rat hippocampal CA1 pyramidal neurons. *J. Physiol.* 543:465–80

Nolan MF, Malleret G, Dudman JT, Buhl DL, Santoro B, et al. 2004. A behavioral role for dendritic integration: HCN1 channels constrain spatial memory and plasticity at inputs to distal dendrites of CA1 pyramidal neurons. *Cell* 119:719–32

Oswald AM, Chacron MJ, Doiron B, Bastian J, Maler L. 2004. Parallel processing of sensory input by bursts and isolated spikes. *J. Neurosci.* 24:4351–62

Oviedo H, Reyes AD. 2002. Boosting of neuronal firing evoked with asynchronous and synchronous inputs to the dendrite. *Nat. Neurosci.* 5:261–66

Pinsky PF, Rinzel J. 1994. Intrinsic and network rhythmogenesis in a reduced Traub model for CA3 neurons. *J. Comput. Neurosci.* 1:39–60

Poirazi P, Brannon T, Mel BW. 2003a. Arithmetic of subthreshold synaptic summation in a model of a CA1 pyramidal cell. *Neuron* 37:977–87

Poirazi P, Brannon T, Mel BW. 2003b. Pyramidal neuron as a 2-layer neural network. *Neuron* 37:989–99

Poirazi P, Mel BW. 2001. Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron* 29:779–96

Polsky A, Mel BW, Schiller J. 2004. Computational subunits in thin dendrites of pyramidal cells. *Nat. Neurosci.* 7:621–27

Rall W. 1964. Theoretical significance of dendritic trees for neuronal input-output relations. In *Neural Theory and Modeling*, ed. R Reiss, pp. 73–97. Stanford, CA: Stanford Univ. Press

Rall W, Burke RE, Smith TG, Nelson PG, Frank K. 1967. Dendritic location of synapses and possible mechanisms for the monosynaptic EPSP in motoneurons. *J. Neurophysiol.* 30:1169–93

Rose CR, Konnerth A. 2001. NMDA receptor-mediated Na^+ signals in spines and dendrites. *J. Neurosci.* 21:4207–14

Rosen MJ. 1972. A theoretical neural integrator. *IEEE Trans. Biomed. Eng.* 19:362–67

Schaefer AT, Larkum ME, Sakmann B, Roth A. 2003. Coincidence detection in pyramidal neurons is tuned by their dendritic branching pattern. *J. Neurophysiol.* 89:3143–54

Schiller J, Major G, Koester HJ, Schiller Y. 2000. NMDA spikes in basal dendrites of cortical pyramidal neurons. *Nature* 404:285–89

Schiller J, Schiller Y. 2001. NMDA receptor-mediated dendritic spikes and coincident signal amplification. *Curr. Opin. Neurobiol.* 11:343–48

Schiller J, Schiller Y, Stuart G, Sakmann B. 1997. Calcium action potentials restricted to distal apical dendrites of rat neocortical pyramidal neurons. *J. Physiol.* 505:605–16

Schwindt PC, Crill WE. 1995. Amplification of synaptic current by persistent sodium conductance in apical dendrite of neocortical neurons. *J. Neurophysiol.* 74:2220–24

Scott EK, Luo L. 2001. How do dendrites take their shape? *Nat. Neurosci.* 4:359–65

Segev I, London M. 1999. A theoretical view of passive and active dendrites. In *Dendrites*, ed. G Stuart, N Spruston, M Häusser, pp. 205–30. Oxford, UK: Oxford Univ. Press

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Segev I, London M. 2000. Untangling dendrites with quantitative models. *Science* 290:744–50

Single S, Borst A. 1998. Dendritic integration and its role in computing image velocity. *Science* 281:1848–50

Smith DJ, Rubel EW. 1979. Organization and development of brain stem auditory nuclei of the chicken: dendritic gradients in nucleus laminaris. *J. Comp. Neurol.* 186:213–39

Sobel E, Tank DW. 1994. In vivo Ca^{2+} dynamics in a cricket auditory neuron: an example of chemical computation. *Science* 263:823–26

Softky W. 1994. Sub-millisecond coincidence detection in active dendritic trees. *Neurosci.* 58:13–41

Somogyi P, Tamas G, Lujan R, Buhl EH. 1998. Salient features of synaptic organisation in the cerebral cortex. *Brain Res. Rev.* 26:113–35

Stuart G, Sakmann B. 1995. Amplification of EPSPs by axosomatic sodium channels in neocortical pyramidal neurons. *Neuron* 15:1065–76

Stuart G, Schiller J, Sakmann B. 1997. Action potential initiation and propagation in rat neocortical pyramidal neurons. *J. Physiol.* 505:617–32

Stuart G, Spruston N, Sakmann B, Häusser M. 1997. Action potential initiation and backpropagation in central neurons. *Trends Neurosci.* 20:125–31

Stuart GJ, Häusser M. 2001. Dendritic coincidence detection of EPSPs and action potentials. *Nat. Neurosci.* 4:63–71

Stuart GJ, Spruston N. 1998. Determinants of voltage attenuation in neocortical pyramidal neuron dendrites. *J. Neurosci.* 18:3501–10

Svoboda K, Helmchen F, Denk W, Tank DW. 1999. Spread of dendritic excitation in layer 2/3 pyramidal neurons in rat barrel cortex in vivo. *Nat. Neurosci.* 2:65–73

Taylor WR, Vaney DI. 2002. Diverse synaptic mechanisms generate direction selectivity in the rabbit retina. *J. Neurosci.* 22:7712–20

Trachtenberg JT, Chen BE, Knott GW, Feng G, Sanes JR, et al. 2002. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. *Nature* 420:788–94

Vetter P, Roth A, Häusser M. 2001. Action potential propagation in dendrites depends on dendritic morphology. *J. Neurophysiol.* 83:3177–82

Wang SS, Denk W, Häusser M. 2000. Coincidence detection in single dendritic spines mediated by calcium release. *Nat. Neurosci.* 3:1266–73

Waters J, Helmchen F. 2004. Boosting of action potential backpropagation by neocortical network activity in vivo. *J. Neurosci.* 24:11127–36

Williams SR. 2004. Spatial compartmentalization and functional impact of conductance in pyramidal neurons. *Nat. Neurosci.* 7:961–67

Williams SR, Stuart GJ. 1999. Mechanisms and consequences of action potential burst firing in rat neocortical pyramidal neurons. *J. Physiol.* 521:467–82

Williams SR, Stuart GJ. 2000. Site independence of EPSP time course is mediated by dendritic I(h) in neocortical pyramidal neurons. *J. Neurophysiol.* 83:3177–82

Williams SR, Stuart GJ. 2002. Dependence of EPSP efficacy on synapse location in neocortical pyramidal neurons. *Science* 295:1907–10

Williams SR, Stuart GJ. 2003. Role of dendritic synapse location in the control of action potential output. *Trends Neurosci.* 26:147–54

Yuste R, Gutnick MJ, Saar D, Delaney KR, Tank DW. 1994. Ca^{2+} accumulations in dendrites of neocortical pyramidal neurons: an apical band and evidence for two functional compartments. *Neuron* 13:23–43

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Annual Review of
Neuroscience

Volume 28, 2005

Contents

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

Genetics of Brain Structure and Intelligence
Arthur W. Toga and Paul M. Thompson 1

The Actin Cytoskeleton: Integrating Form and Function at the Synapse
Christian Dillon and Yukiko Goda 25

Molecular Pathophysiology of Parkinson's Disease
Darren J. Moore, Andrew B. West, Valina L. Dawson, and Ted M. Dawson 57

Large-Scale Genomic Approaches to Brain Development and Circuitry
Mary E. Hatten and Nathaniel Heintz 89

Autism: A Window Onto the Development of the Social and the Analytic Brain
Simon Baron-Cohen and Matthew K. Belmonte 109

Axon Retraction and Degeneration in Development and Disease
Liqun Luo and Dennis D.M. O'Leary 127

Structure and Function of Visual Area MT
Richard T. Born and David C. Bradley 157

Growth and Survival Signals Controlling Sympathetic Nervous System Development
Natalia O. Glebova and David D. Ginty 191

Adult Neurogenesis in the Mammalian Central Nervous System
Guo-li Ming and Hongjun Song 223

Mechanisms of Vertebrate Synaptogenesis
Clarissa L. Waites, Ann Marie Craig, and Craig C. Garner 251

Olfactory Memory Formation in *Drosophila*: From Molecular to Systems Neuroscience
Ronald L. Davis 275

The Circuitry of V1 and V2: Integration of Color, Form, and Motion
Lawrence C. Sincich and Jonathan C. Horton 303

Molecular Gradients and Development of Retinotopic Maps
Todd McLaughlin and Dennis D.M. O'Leary 327

Neural Network Dynamics
Tim P. Vogels, Kanaka Rajan, and L.F. Abbott 357

The Plastic Human Brain Cortex
Alvaro Pascual-Leone, Amir Amedi, Felipe Fregni, and Lotfi B. Merabet 377

An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance
Gary Aston-Jones and Jonathan D. Cohen 403

Neuronal Substrates of Complex Behaviors in *C. elegans*
Mario de Bono and Andres Villu Maricq 451

Dendritic Computation
Michael London and Michael Häusser 503

Optical Imaging and Control of Genetically Designated Neurons in Functioning Circuits
Gero Miesenböck and Ioannis G. Kevrekidis 533

INDEXES

Subject Index 565

Cumulative Index of Contributing Authors, Volumes 19–28 577

Cumulative Index of Chapter Titles, Volumes 19–28 582

ERRATA

An online log of corrections to *Annual Review of Neuroscience* chapters may be found at http://neuro.annualreviews.org/

Annu. Rev. Neurosci. 2005.28:503-532. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF WASHINGTON on 09/07/05. For personal use only.

DENDRITIC INTEGRATION OF EXCITATORY SYNAPTIC INPUT

Jeffrey C. Magee

A fundamental function of nerve cells is the transformation of incoming synaptic information into specific patterns of action potential output. An important component of this transformation is synaptic integration — the combination of voltage deflections produced by a myriad of synaptic inputs into a singular change in membrane potential. There are three basic elements involved in integration: the amplitude of the unitary postsynaptic potential, the manner in which non-simultaneous unitary events add in time (temporal summation), and the addition of unitary events occurring simultaneously in separate regions of the dendritic arbor (spatial summation). This review discusses how passive and active dendritic properties, and the functional characteristics of the synapse, shape these three elements of synaptic integration.

CABLE FILTERING
Dendrites have been commonly modelled as cables and the flow of current between two points of a dendrite has been usually assumed to decay as a result of filtering along the process.

HEBBIAN PROCESSES
Plastic processes that require temporal coincidence between incoming synaptic activity and postsynaptic depolarization.

AXIAL RESISTANCE
The resistance to the flow of ionic current along an axon or a dendrite. Axial resistance decreases as a function of the radius of the process and increases as a function of its length.

*Neuroscience Center,
Louisiana State University
Medical Center, 2,020
Gravier Street, New Orleans,
Louisiana 70112, USA.
e-mail: jmagee@lsumc.edu*

Most neurons of the mammalian central nervous system receive thousands of excitatory and inhibitory synaptic inputs that are widely spread across their intricate dendritic arbors. Whereas synaptic input is broadly distributed, action potential output usually occurs in a more localized region of the proximal axon[1,2]. Because of this spatial arrangement, the distance between the various synaptic inputs and the final integration site can vary to a great degree. The combination of this large variation in synaptic distance and the CABLE FILTERING properties of dendrites can, in theory, cause the amplitude and temporal characteristics of functionally similar inputs to be highly variable at the final integration site[3-8].

The impact of any given synapse on the neuron firing will depend primarily on the location of the synapse along the dendrite, unless the filtering properties of the dendrites are compensated for by specific mechanisms. This location dependence could be so profound that even the basic elements of synaptic integration would depend on the pattern of spatial input. Furthermore, the distorting effects of the dendritic arbor could prevent a neuron from determining accurately the total incoming synaptic activity, decrease the effectiveness of HEBBIAN PROCESSES and increase the variability of action potential discharge. In consequence, the existence of mechanisms for reducing the dependence of synaptic effectiveness on input location could have important advantages for neuronal function.

Although theoretical analyses have predicted a clear location-dependent variability of synaptic input, many functional properties of the neuron can actually reduce this variability. They include the passive and active properties of the postsynaptic cell membrane, as well as various pre- and postsynaptic properties of the synapses themselves. In this review, I provide a brief description of the filtering effects of dendrites and how they affect propagation of the excitatory postsynaptic potential (EPSP). Next, I discuss data indicating that, in spite of these filtering effects, the impact of a given synapse may not be strictly dependent on the location of the synapse. The mechanisms believed to be responsible for the reduction of location dependence are examined, before I consider finally the potential consequences of location-independent synaptic input on neuronal function.

Dendritic filtering

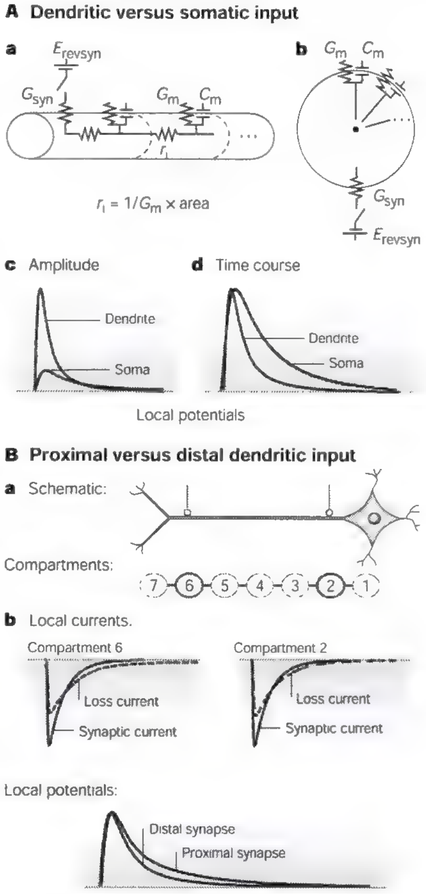
The filtering effects of dendritic arbors have been appreciated for decades. About 40 years ago, Wilfrid Rall began an influential series of studies modelling the passive electrical properties of dendrites using three cable parameters: AXIAL RESISTANCE (r_i), MEMBRANE CAPACITANCE (C_m) and MEMBRANE CONDUCTANCE (G_m)[3-5,7]. The flow of

© 2000 Macmillan Magazines Ltd

Box 1 | Theoretical dependence of synaptic input on location

Cell bodies and neuronal processes are commonly modelled as electrical circuits, models that take into account the geometry of the structure being analysed and that make strong predictions about the dependence of current flow on shape. The top panel of the figure shows the equivalent circuit for dendritic cables (Aa) and somatic spheres (Ab) where G_m and C_m are the resting conductance and capacitance of a unit of membrane and r_i is the axial resistance of the intracellular side of a unit of cable. A synapse is represented as a synaptic conductance (G_{syn}) in series with the synaptic battery (E_{revsyn}). In the case of a dendrite, synaptic current will flow through the membrane into the parallel G_m and C_m as well as into the next compartment through r_i. In contrast, the spherical soma does not have extra pathways for current flow and all synaptic current will flow into the somatic G_m and C_m. Similarly, geometry also affects input resistance such that it is inversely proportional to compartment area (Ac). As a result, the local amplitude of an excitatory postsynaptic potential (EPSP) generated by the same synaptic conductance will be larger in the small-diameter, high-resistance dendrite than in the spherical soma. At the same time, the availability of pathways for current flow other than G_m and C_m causes synaptic current to flow away from the synaptic compartment (Ad), speeding up the decay of the dendritic EPSP in relation to the soma.

How does this effect of geometry affect synaptic current at different points of the dendritic arbour, such as those schematized in Ba? If the synaptic current (Bb; solid line) is recorded locally in either compartments 6 (left) or 2 (right), it will be identical in both sites. In contrast, the current leaving the synaptic compartment (loss current; dashed line) will not. This observation results from the fact that the proximal region of the dendrite becomes isopotential with the soma much faster than the distal region, reducing the driving force for current flow away from the input site. Because the loss of current is more prolonged in the distal regions, the distal EPSPs are more transient than the proximal ones (Bc) (see REF. 4).



synaptic current into G_m and C_m produces a change in the local membrane potential that depends on the neuronal geometry and on these three cable parameters. The axial resistance causes a continuous voltage drop between the site of input and the soma. As a result, the driving force for current flow across G_m decreases progressively with distance from the cell body. Furthermore, the temporary storage of synaptic charge in C_m slows the kinetic properties of the resulting change in membrane potential (BOX 1).

Because r_i increases as a function of length and C_m increases as a function of membrane area in any given dendrite, distal synapses will experience more amplitude and kinetic filtering than proximal synapses. As a result, distal EPSPs should have smaller amplitudes than proximal synaptic events. In contrast, the increased filtering experienced by distal inputs should have a different effect on the summation of several EPSPs. Dendritic filtering should prolong the time window over which activity at the slower, distal inputs could overlap with voltage changes caused by other synapses, thereby extending the period over which summation of distal synapses can effectively occur. Therefore, the increased dendritic filtering experienced by distal synaptic inputs should have two distinct effects on their integration properties: it should produce EPSPs of smaller amplitude but, at the same time, longer periods of temporal integration at the soma[3–8] (FIG. 1).

Rall's theoretical analysis predicted that synapses could be functionally distinguished on the basis of dendritic location and that the impact of individual synapses would decrease with distance from the soma. For Rall, the slower, distal EPSP would only cause a subthreshold change in somatic membrane potential, whereas activity from a few of the larger, more transient, proximal synapses would produce sharp, spike-triggering depolarizations[3]. In modern terms, the neuron would operate in a TEMPORAL-INTEGRATION MODE for the distal inputs and in a COINCIDENCE-DETECTION MODE for the proximal inputs[9]. The cable properties of dendrites have therefore been thought to exert a shaping force on the EPSP that, if not countered, could have a important impact on the temporal pattern of action potential output produced by any given spatio-temporal pattern of synaptic input.

MEMBRANE CAPACITANCE
The cell membrane separates and stores electrical charge, thereby producing a relatively large electrical capacitance, which increases as a function of membrane area.

MEMBRANE CONDUCTANCE
The sum of all of the ionic conductances of the membrane. It is the reciprocal of the membrane resistance.

TEMPORAL-INTEGRATION MODE
Firing of action potentials in response to the summation of temporally dispersed synaptic activity. The synaptic input at the soma has small amplitude and long duration.

COINCIDENCE-DETECTION MODE
In this mode, a cell fires action potentials in response to the simultaneous arrival of a small number of usually large amplitude, short duration inputs.

© 2000 Macmillan Magazines Ltd

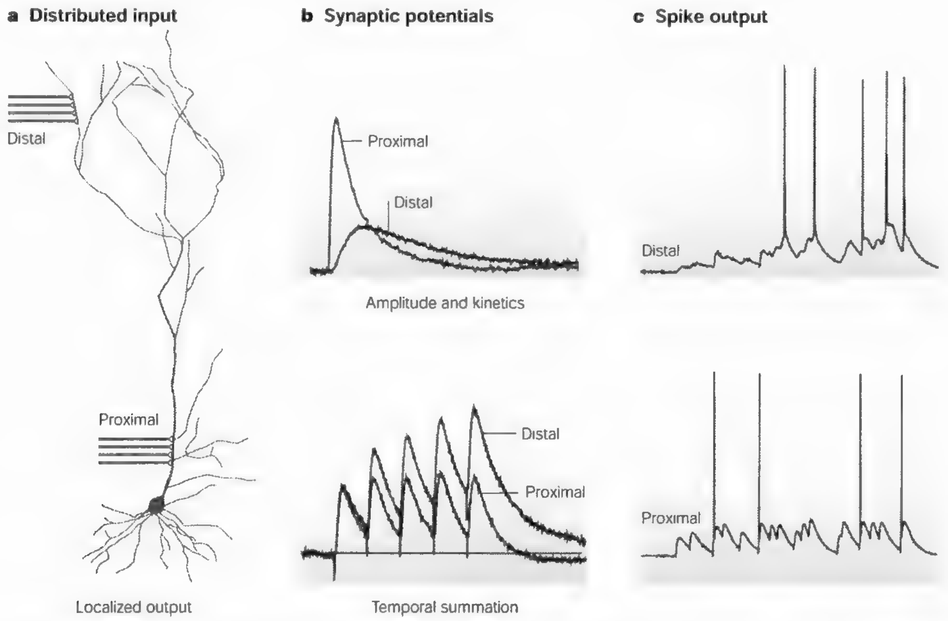


Figure 1 | **The filtering effects of dendrites should be accompanied by location-dependent synaptic integration.** **a** | A CA1 pyramidal neuron with synaptic input distributed across the dendritic arbor. **b** | Dendritic filtering causes distal inputs to be smaller and slower at the soma. The temporal summation is greater for slower distal events than for the proximal input. **c** | When compared with proximal inputs, the smaller, slower distal inputs should produce a very different action potential output pattern for the same pattern of input. All traces represent somatic membrane potential in response to electrical stimulation (**b**) or excitatory postsynaptic current (EPSC)-shaped current injections (**c**) in the presence of 50 µm ZD7288, a blocker of transient outward current, carried by dendritic hyperpolarization-activated potassium channels (I_h).

A series of recent experiments has tested the validity of those theoretical predictions. As I discuss in the following section, most of the evidence obtained so far indicates that, contrary to the theory, synaptic integration is largely independent of input location.

Location dependence of synaptic integration

The location dependence of synaptic input has been examined in various cell types, especially in hippocampal CA1 and neocortical pyramidal cells, as well as in spinal motor neurons. In contrast to Rall's predictions, there is now considerable evidence to indicate that, within a single class of inputs, most components of synaptic integration may show minimal dependence on the location of the synapse. This evidence is now examined with an emphasis on data from Schaffer collateral synapses on hippocampal CA1 pyramidal neurons.

Unitary EPSP amplitude. In the first examination of the location dependence of synaptic input to CA1 neurons, Andersen *et al.*[10] concluded that proximal and distal Schaffer collateral synapses could elicit similar patterns of action potential firing and found that the shape of the EPSP was relatively independent of synapse location. Subsequently, Redman and colleagues reached a similar conclusion through a statistical analysis of excitatory postsynaptic current (EPSC) fluctuations. This group obtained evidence that the amplitude of the QUANTAL CURRENT at the soma was independent of the location of the synapse along the dendrite[11]. Nevertheless, previous work in his own laboratory had shown that evoked EPSP amplitude seemed to decrease with distance from the soma because the amplitude of EPSPs evoked by stimulation of single CA3 neurons decreased as the time to reach the peak amplitude increased[12]. To reconcile the data, Redman and colleagues proposed that the QUANTAL CONTENT of evoked Schaffer collateral EPSP decreased with distance from the soma[11]. Subsequent work has shown that the quantal EPSP amplitude of single sensory Ia fibres in spinal motor neurons is location independent[13,14]. More recently, a study using localized release of CAGED GLUTAMATE has shown that the amplitude of the evoked current measured at the soma is independent of the site of glutamate application along the apical dendrite of hippocampal CA1 pyramidal neurons[15]. Similar results have also been reported for pyramidal cells of layer V of the neocortex[16].

The ability to record synaptic activity from several locations on the same neuron simultaneously has provided the opportunity to test directly many long-standing ideas about the dendritic processing of synaptic information (FIG. 2). Using dual whole-cell patch-clamp recordings in combination with localized MINIMAL STIMULATION TECHNIQUES, it has been found that the amplitude of unitary EPSP at the soma of CA1 pyramidal cells is

QUANTAL CURRENT
The current elicited by the transmitter released from a single synaptic vesicle.

QUANTAL CONTENT
The number of quanta released per action potential.

CAGED GLUTAMATE
An inactive derivative of glutamate that can be transformed into the active transmitter, usually by photolysis of the precursor. It provides an efficient means for the spatially restricted application of glutamate.

MINIMAL STIMULATION TECHNIQUES
Methods to elicit transmitter release from a few (ideally one) synaptic contacts.

© 2000 Macmillan Magazines Ltd

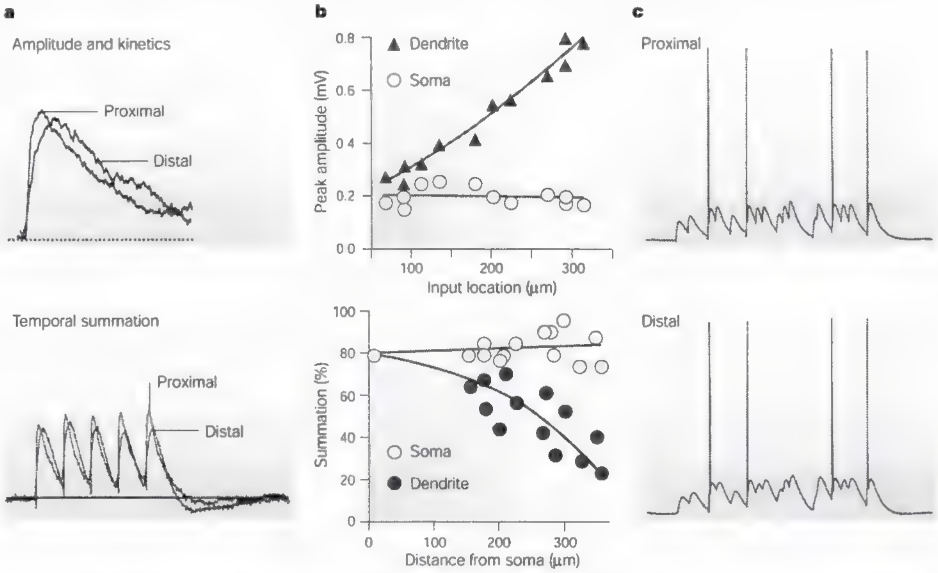


Figure 2 | **Synaptic integration in CA1 pyramidal neurons is independent of location.** **a** | Upper traces: The average unitary excitatory postsynaptic potentials (EPSPs) recorded at the soma of a neuron receiving distal and proximal input. Somatic EPSP amplitude is similar in spite of location differences. Lower traces: The amount of temporal summation at the soma is the same for a 50 Hz train of stimuli applied to distal (~300 μm) or to proximal Schaffer collateral inputs (~50 μm). **b** | Upper plot: Mean EPSP amplitude as a function of input distance from the cell body for recordings made at the soma (yellow circles) or at the dendrite (red triangles). EPSP amplitude at the site of input increases with distance from the soma, whereas somatic amplitude stays relatively constant. Lower plot: temporal summation during a 50 Hz train of stimuli as a function of distance from the cell body for recordings made at the soma (yellow circles) or the dendrite (red circles). The plot shows that summation at the site of input decreases with distance from the soma, whereas somatic summation stays relatively constant. **c** | Somatic voltage in response to current injected through the somatic electrode (proximal) or through the dendritic pipette located about 350 μm away from the soma (distal). The temporal pattern of action potential output is independent of input location. (Figure adapted with permission from *Nature Neurosci.* (REFS 17, 19) © (2000) Macmillan Magazines Ltd.)

independent of synapse location across the entire range of Schaffer collateral inputs[17]. These recordings have also shown that the EPSP amplitude recorded precisely at the dendritic site of input experiences a progressive threefold to fourfold increase as a function of distance from the soma. In other words, if the amplitude of an input is expected to decrease by 20% while propagating to the soma (from 1.0 to 0.8 mV), then the amplitude recorded at the synaptic site will be 25% larger (from 1.0 to 1.25 mV)[17]. If the input is to decay by 70% (from 1.0 to 0.3 mV), then a threefold increase in local amplitude is observed (from 1.0 to 3 mV)[17]. These data have shown that local EPSP amplitude is increased just enough to counterbalance the filtering effects of the dendrite such that all inputs have roughly the same amplitude after propagation to the soma (FIG. 2)[17].

EPSP kinetics and temporal summation. Similar recordings in CA1 pyramidal neurons have shown that, although the rise times of somatic EPSPs show location dependence and increase with distance from the soma, the widths of the EPSPs do not (FIG. 2)[17–21]. As all EPSPs last roughly the same time in these neurons, temporal summation of several EPSPs at the soma is virtually independent of synapse location for a large range of stimulation frequencies (5–100 Hz) and amplitudes (0.5–5 mV)[19]. A similar location independence of somatic EPSP HALF-WIDTHS and of temporal summation has also been reported for neocortical layer V pyramidal neurons[21].

How can the more filtered, distal EPSPs have the same duration as proximal events? As was observed for the normalization of EPSP amplitude, the local dendritic EPSP half-widths and DECAY TIME CONSTANTS progressively decrease with distance from the soma. Again, the local EPSP shape is modified just enough to counterbalance the filtering effects of the dendrite and to greatly reduce the location dependence of EPSP duration and temporal summation[19,21].

Spatial summation. In contrast to unitary EPSP amplitude and temporal summation, a small degree of location dependence has been observed for one aspect of spatial summation in CA1 pyramidal neurons. Using iontophoretically applied glutamate, Cash and Yuste[22] have shown that the summation of two stimuli applied to different points on the apical trunk of the dendrite is smaller than the expected linear sum of the individual events. This sublinearity was evident only for distant inputs and disappeared with repetitive stimulation. Furthermore, the location dependence was seen only for inputs to the apical trunk and not to other parts of the dendritic arbor. Also, the degree of this sublinearity is

EPSP HALF-WIDTH
The duration of an EPSP at the point at which its amplitude is half of the peak value.

DECAY TIME CONSTANTS
The initial decay of an EPSP can usually be fitted by a single-exponential function. The time constant derived from this fit describes how quickly an EPSP decays.

© 2000 Macmillan Magazines Ltd

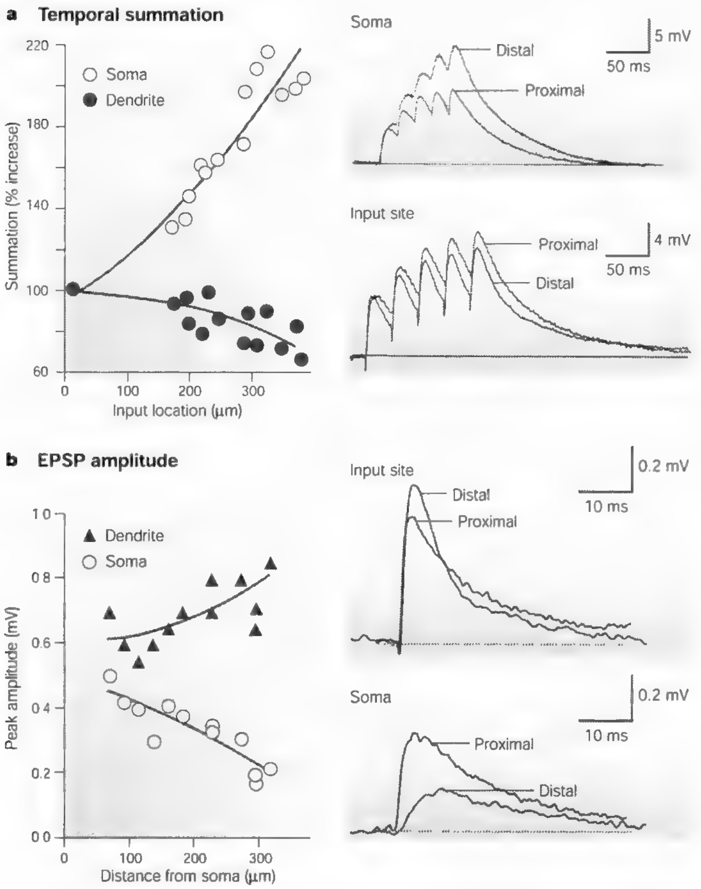


Figure 3 | **EPSC-shaped current injections show that the passive morphology of CA1 pyramidal neurons does not fully counter location-dependent synaptic variability. a** | Temporal summation for a train of 50 Hz excitatory postsynaptic current (EPSC)-shaped current injections measured at the soma or at the dendrite in the presence of ZD7288, a blocker of transient outward current, carried by dendritic hyperpolarization-activated potassium channels (I_h). Although the local summation decreases as the input site is moved away from the soma (red symbols and top traces), temporal summation of the propagated excitatory postsynaptic potentials (EPSP) at the soma remains largely dependent on the input location (yellow symbols and bottom traces). This indicates that the passive membrane properties of the cell are not capable of normalizing temporal summation at the soma. **b** | EPSP amplitude at the site of input (red triangles, top traces) or at the soma (yellow circles and bottom traces) for EPSC-shaped current injections. Although the local EPSP amplitude increases with distance, the somatic EPSP amplitude continues to show a large amount of location dependence, indicating that postsynaptic membrane properties cannot normalize unitary EPSP amplitude. (Figure adapted with permission from *Nature Neurosci.* (REFS 17, 19) © (2000) Macmillan Magazines Ltd.)

likely to be voltage dependent, as the summation of smaller unitary depolarizations seems to behave more linearly and to show a reduced location dependence[22] (J.C.M., unpublished observations). Although it is still difficult to draw definitive conclusions from the available data, there does seem to be a small level of location dependence to some aspects of spatial summation in CA1 pyramidal neurons.

In summary, evidence from different cell types is beginning to show that, contrary to theoretical expectations, the three components of synaptic integration — unitary amplitude, and temporal and spatial summation — show minimal location dependence. What dendritic and synaptic mechanisms are responsible for determining the integration properties and for reducing the variability of the synaptic responses along the apical dendrite in these cells? The different mechanisms can be divided into three general categories: morphological (passive properties of the membrane), active properties of the membrane and synaptic mechanisms. The next sections discuss evidence for the effect of each of these possible mechanisms on the separate components of synaptic integration.

ISOPOTENTIALITY
The degree to which the dendritic compartments are all at the same potential.

Neuron morphology

The morphology of the neuron could affect the amplitude and the shape of local EPSPs and potentially counter some aspects of the location dependence of synaptic variability. It has been shown both in multicompartmental computer models and in dendritic recordings that the local input resistance for a transient event increases, whereas local capacitance decreases, with distance from the soma[17,23–25]. This effect has generally been ascribed to the decrease in compartment size that occurs as the diameter of the distal dendrites decreases progressively[23,24]. The subsequent increase in total r_i with distance from the soma can also contribute to this effect.

Morphology also determines the ISOPOTENTIALITY of the local dendritic compartments. Dendritic arbors are complex branched structures in which current flows away from and towards the synaptic-input site, and not only into the membrane resistance and capacitance[23–25]. The specific direction of current flow will be determined by the potential differences established between the different regions of the cell. So, the other paths for current flow provided by the dendritic branches dictate that dendrites are less isopotential than the large spherical soma[8,23,24]. What is the impact of these differences on synaptic integration?

Unitary EPSP amplitude. The changes in input resistance and capacitance that occur in distal dendritic compartments cause EPSP amplitude at the site of input to increase with distance from the soma, as has been observed on CA1 pyramidal cells[17,23–25]. Although the increase in depolarization is not necessarily proportional to the increase in input resistance, the increase in local EPSP amplitude is substantial and depends on passive parameters as well as on neuron morphology[17,23,24]. In dendritic recordings, EPSC-shaped current injections cause voltage deflections that also increase in amplitude with distance from the soma (~50% by 325 μm; FIG. 3)[17]. The amount of increase, however, is notably smaller than that observed in theoretical models and, more importantly, it is much smaller than that shown for the actual unitary synaptic input. Accordingly, and in contrast to that observed for synaptic inputs, these EPSC-shaped current injections into the dendrite show a sharp location dependence, as their amplitude decreases for the more distal injection sites (FIG. 3)[17,23,24]. Therefore, it seems that the increase in local dendritic input resistance

© 2000 Macmillan Magazines Ltd

contributes only slightly to the increases in local EPSP amplitude that are responsible for normalizing the EPSP amplitude at the soma[23–25].

EPSP kinetics and temporal summation. The changes in compartment size and branching that occur with distance from the soma can also shape the kinetics of the local EPSPs. As mentioned above, the spread of current throughout the dendrites is sufficient to cause an initially faster decay of synaptic depolarization in the dendritic compartment than in the soma. This effect is related to the proximity of the synaptic input to the cell body; more distal inputs show the fastest initial decay phase[4,17,21,23–25]. However, once the cell becomes isopotential, the decay of the EPSP would be the same at every point.

In conclusion, as has been observed for the local EPSP amplitude, EPSP half-width and local temporal summation do decrease progressively with distance from the soma as a consequence of neuron morphology (FIG. 3)[4,17,23,24]. However, as was also the case for the EPSP amplitude, the impact of the passive structure on the time course and summation of the synaptic potential is not adequate to completely counteract the filtering effects of the dendrite[17,21]. Therefore, other mechanisms besides the passive properties of the dendrites are required to reduce the effect of filtering by an extra factor of two or three. So, whereas the dendritic morphology of pyramidal neurons can contribute to spatial gradients that affect synaptic integration, it is unlikely that passive geometry alone can substantially reduce the location-dependent variability of synaptic inputs[23–25]. Although this idea has been confirmed in neurons such as CA1 and neocortical pyramidal cells, the possibility remains that the different morphology of other cell types, such as dentate granule cells, CA3 pyramidal neurons and CA3 interneurons, could be more important in normalizing EPSP amplitude[25].

Spatial summation. The larger, local EPSPs produced by the higher input resistance of the distal dendrites will experience reduced driving forces for synaptic current flow compared with more proximal synaptic events[22–24]. This effect could potentially induce a moderate location dependence of the EPSP at the soma during the spatial summation of inputs. In this case, a greater sublinear summation should be seen for distal versus proximal inputs and, indeed, such location dependence has been observed for distal inputs to the apical trunk of CA1 pyramidal neurons[22]. However, the sublinearity of this input was smaller than expected and was not present for other input locations. These results, along with pharmacological data that will be discussed below, have led to the belief that the impact of the passive structure of the dendrites is overwhelmed by other mechanisms such as the opening of dendritic voltage-gated ion channels[22]. In fact, this conclusion is representative of the impact of pyramidal cell morphology on most aspects of synaptic integration. Whereas the passive structure of the dendrites does indeed have an important impact on integration, this impact is limited in comparison with other mechanisms in the dendrites, as will be discussed in the next section.

Active properties of dendrites

Voltage-gated ion channels have been found in the dendrites of virtually every mammalian neuron studied so far, although their exact subcellular distribution is specific for each cell type (for a review, see REF. 26). The currents carried by these channels are important in shaping the integrative properties of neurons. For instance, the inward currents through Na^+ and Ca^{2+} channels could provide extra ion flux to replace the synaptic current lost across the membrane resistance. The combination of an initial inward current followed by a transient outward current, carried by dendritic hyperpolarization-activated K^+ channels (I_h), could also reduce the filtering effects of the dendritic capacitance by charging and discharging the membrane during the EPSP[27–29]. Indeed, there have been many reports on the capability of voltage-gated ion channels to shape both single and repetitive excitatory synaptic input in hippocampal and neocortical pyramidal neurons[28–37]. Furthermore, the voltage-gated ion channels present in dendrites allow action potentials initiated at the axon to propagate back into the dendrites, where they can then sum with incoming synaptic activity[2,26]. This summation tends to increase the amplitude of the back propagating spikes and has been reported to lead to the generation of Ca^{2+}-mediated potentials in some cells[38,39]. This dendritic electrogenesis can greatly enhance the impact of very distal synaptic input such as that received by the apical tufts of layer V neocortical pyramidal cells[39,40]. In summary, the function of active channels in shaping EPSP integration will depend on the precise channel distribution and on the neuronal geometry of any particular cell. Nevertheless, some general principles have emerged for pyramidal neurons, as discussed below.

Unitary EPSP amplitude. In CA1 neurons[26] and in neocortical pyramidal cells[28,41], the impact of the outward currents tends to increase with distance from the soma because the density of transient K^+ and I_h channels increases along the length of the apical dendrites. As a consequence, the main impact of dendritic ion channels on unitary synaptic input should not be a boosting of the distal synaptic amplitude. Instead, the channels should increase the location dependence of somatic EPSP amplitude by reducing distal EPSPs even more than proximal ones. However, for these channels to actually participate in shaping synaptic inputs, the amplitude of the EPSPs needs to be sufficient to activate a substantial fraction of the channels in question. For instance, assuming a resting potential of –70 mV, the depolarization should be roughly 10–15 mV for the activation of Na^+ and low-voltage-activated Ca^{2+} channels, and 15–20 mV for the opening of A-type K^+ channels[26]. I_h is active at rest but the deactivation rate of these channels is so slow that current through this channel population will mostly affect only the EPSP decay and not the amplitude[18,19,21]. Therefore, although the exact effect of voltage-gated ion channels on unitary EPSP amplitude has yet to be determined, their effect on these small amplitude events should be limited[17]. On the other hand, summed EPSPs will be larger in amplitude and therefore shaped to a larger degree by dendritic ion channels.

© 2000 Macmillan Magazines Ltd

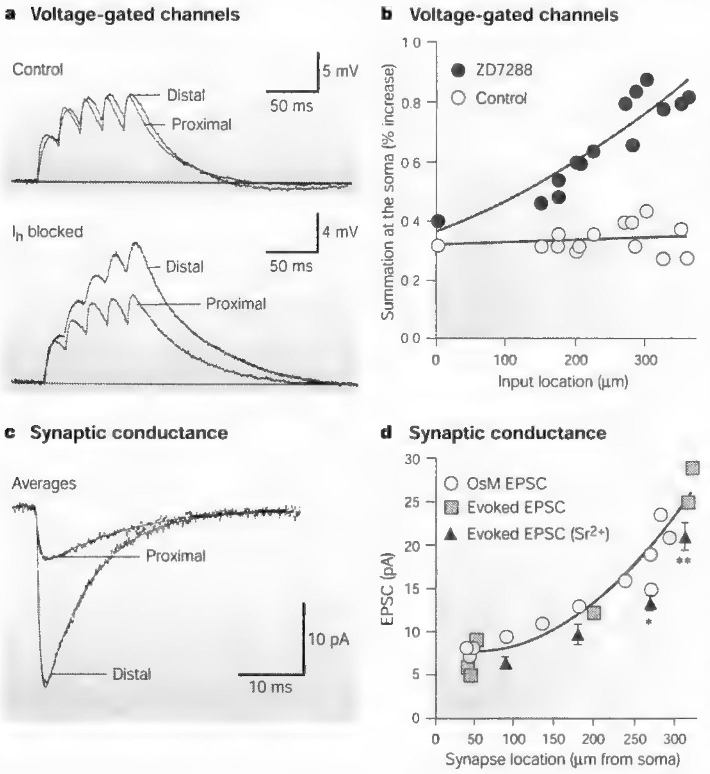


Figure 4 | **Other mechanisms such as active membrane and synaptic properties are required to completely counter location-dependent synaptic variability. a** | Membrane voltage recorded at the soma in response to a train of five excitatory postsynaptic current (EPSC)-shaped current injections delivered at 50 Hz through the dendritic or the somatic pipette. Upper traces were obtained under control conditions and lower traces were obtained under blockade of transient outward current, carried by dendritic hyperpolarization-activated potassium channels (I_h) (20 μM ZD7288). Summation at the soma is larger for dendritic current injections but it does not change notably for somatic injections. **b** | Excitatory postsynaptic potentials (EPSP) summation at the soma plotted as a function of the dendritic site of current injection under control conditions (yellow circles) or in the presence of ZD7288 (red circles). When I_h is blocked, the amount of summation at the soma increases as input location is moved away. **c** | Average of EPSC, evoked by the localized application of a solution of high osmolarity, and recorded from a proximal (~100 μm) and a distal (~280 μm) site in the apical dendrite of two different neurons. **d** | Plot of mean EPSC amplitude as a function of input distance from the soma for EPSCs evoked by high osmolarity (yellow circles) or by electrical stimulation in normal (blue squares) or Sr^{2+}-containing extracellular solution (red triangles). Sr^{2+} causes desynchronized release of transmitter and provides a better estimate of the amplitude of the unitary synaptic response. (Figure adapted with permission from *Nature Neurosci.* (REFS 17, 19) © (2000) Macmillan Magazines Ltd.)

EPSP kinetics and temporal summation. The impact of dendritic channels on EPSP shape and summation has been studied in pyramidal cells in some detail. For short time intervals (0–5 ms), the summation of two inputs closely spaced in time seems to be linear or slightly supralinear[22,34,35,37], whereas EPSP summation over longer intervals (5–100 ms) has been reported to be sublinear[19,35,37]. This nonlinear behaviour is affected by blockade of voltage-dependent channels. For example, blockade of Na^+, Ca^{2+} or NMDA receptor channels decreases or abolishes the supralinearity, whereas I_h and K^+ channel blockade eliminates the sublinear summation[21,22,34,35,37]. Large synchronous synaptic activity (>30 mV) can initiate dendritic Na^+ spikes that usually do not propagate as action potentials to the soma[19,42]. The ability of a given synaptic input to sum supralinearly is enhanced by repetitive activity, so that large EPSPs occurring later in a train have an enhanced amplitude and are more likely to initiate dendritic spikes[19,22,27,42]. In CA1 pyramidal neurons, this is presumably related to the inactivation of voltage-gated channels (probably A-type K^+ channels) during the train of activity. Therefore, it seems that although dendritic voltage-gated channels may have a limited impact on the amplitude of unitary synaptic input, they actively shape repetitive synaptic potentials of larger amplitude.

Furthermore, the increased density of I_h channels, found in both CA1 and neocortical pyramidal cells, is prominent in reducing the location dependence of the EPSP half-width and temporal summation[18,19,21,28,32]. The effectively outward active current generated by I_h channel deactivation during synaptic activity produces a hyperpolarization that tends to shorten the duration of local EPSPs and to counter any local temporal summation during repetitive stimuli[19,21]. As the channel density increases with distance from the soma, the size of the outward current generated at the site of the input also increases, producing the spatial gradient of local EPSP duration and summation described above. Because selective blockade of I_h channels causes a profound location-dependent synaptic variability, the current generated by I_h channel deactivation is presumably the most important factor in normalizing temporal summation in CA1 pyramidal neurons (FIG. 4a, b)[19,21].

Spatial summation. A-type K^+ channels present in dendrites seem to be important in shaping spatial EPSP summation in hippocampal pyramidal neurons[22,27,33,35,37]. These channels are prominent in the distal dendritic regions of CA1 and perhaps of other pyramidal cells, where they seem to limit the summation of closely timed distal inputs[22,35,37]. In fact, the location dependence of the sublinear summation described above for CA1 cells is sensitive both to the A-type K^+-channel blocker 4-aminopyridine and to stimuli that inactivate this class of channel[22]. In CA1 cells, blockade of Na^+ and Ca^{2+} channels causes sublinear spatial summation in radial oblique dendrites, whereas only tetrodotoxin, a blocker of Na^+ channels, has the same effect on input to basal dendrites. These data indicate that, at least in CA1 cells, there is an interaction between the active conductances present in the dendrites, which linearizes spatial summation in most of the proximal dendritic arbor[22,33]. In CA3 pyramidal neurons, the summation of distal and proximal input is sublinear even under control conditions. This sublinearity is removed by blocking A-type K^+-channels, but is not affected by Na^+ or Ca^{2+} channel blockade.

Finally, the integration of synaptic activity with back propagating action potentials provides a unique form of interaction between synaptic inputs at different sites of the dendritic arbor. As reported for layer V neocortical pyramidal cells, any distal input arriving within 10 ms

© 2000 Macmillan Magazines Ltd

of other suprathreshold inputs will receive a pronounced boost from the further activation of distal dendritic voltage-gated Na^+ and Ca^{2+} channels[39]. Therefore, although the direct interaction between the widely spaced EPSPs is small, the action potential produced by suprathreshold proximal inputs will effectively sum with the incoming distal input to produce a nonlinear form of integration. Furthermore, this interaction should also show some frequency dependence because the probability of initiating Ca^{2+}-mediated potentials in the distal apical dendrite increases with repetitive activity for a critical range of frequencies that depends on the cell (60–200 Hz)[39,40]. So, it seems that the active properties of dendrites endow neurons with the ability to produce both linear (spatially normalized summation) and highly nonlinear interactions (local dendritic spiking) between synaptic inputs. The exact effect will depend on the specific spatio-temporal characteristics of the input.

Synaptic properties

By modifying the size and the kinetic properties of the synaptic conductance, the synapse itself can equip the neuron with a plethora of potential mechanisms to shape synaptic integration. These mechanisms can be presynaptic (such as the vesicular content of transmitter, the number of vesicles released per site, or the number of release sites per terminal) or postsynaptic (receptor number, density, affinity or single-channel conductance and kinetics).

Unitary EPSP amplitude. As discussed above, the morphology and active properties of CA1 pyramidal neurons are not likely to confer the level of location independence that has been observed for unitary EPSP amplitude in these cells. Therefore, the dendritic mechanisms for reducing location-dependent amplitude variability are likely to be present at the synapse itself. In support of this idea, the local EPSC amplitude in CA1 pyramidal neurons measured under whole-cell voltage clamp at the dendritic site of input was observed to increase roughly threefold across the range of the Schaffer collateral input[11,17] (FIG. 4c, d). The voltage-clamp of these synapses is good but imperfect (>90% voltage attenuation)[17]. Therefore, the passive properties of the dendrites can have a small impact on the measured EPSC amplitude and kinetics. Because of this, the TIME INTEGRAL of the EPSCs has been used as a more accurate filter-independent measure of synaptic conductance[8,23] and has been found to increase more than twofold across the same range of inputs[17]. These data indicate that unitary synaptic conductance increases with distance from the soma in CA1 pyramidal neurons[11,17] and that this increase is the primary mechanism for reducing the location dependence of unitary EPSP amplitude.

The exact synaptic modification responsible for this increase in conductance is unknown at present. Analysis of EPSC rise times and amplitude distributions indicate that the most plausible candidates might be increases in the number of release sites per terminal and/or increases in the postsynaptic area and number of glutamate receptors[17,43–49]. Analogous changes in these synaptic parameters have been previously observed in goldfish MAUTHNER CELL dendrites, where the number of release sites per bouton as well as the total area of the active zones are increased in the more distal dendrites[49]. These data are also consistent with the increased sensitivity to glutamate observed in the distal apical dendrites of both neocortical and hippocampal pyramidal neurons[15,16]. So, it seems that several cell types have used an increase in local synaptic conductance to counterbalance the amplitude filtering effects of their dendrites.

The time course of a synaptic event can also affect the level of dendritic filtering notably and, as a consequence, the degree of location dependence of the synaptic input. In the case of glutamate-mediated transmission, α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor-mediated events have kinetic properties about one order of magnitude faster than NMDA (*N*-methyl-D-aspartate) receptor-mediated EPSPs[50]. Because the amount of charge stored in the dendritic capacitance is directly proportional to the rate of change of the membrane potential, the amount of filtering done by a dendritic arbor will depend on the kinetic features of the potential change[8,23,24]. As a result, the faster AMPA-mediated EPSP will experience a greater amount of amplitude attenuation when compared with the slower NMDA-mediated events. Therefore, distal synapses could increase their impact by raising their NMDA-to-AMPA receptor ratio and, in fact, there is some evidence that perforant path synapses in the most distal regions of CA1 pyramidal neurons may use this strategy[51].

EPSP kinetics, temporal and spatial summation. There are some clear examples of how receptor kinetics affect EPSC duration and temporal integration in hippocampal dentate gyrus basket cells and in neurons from brainstem auditory nuclei[52,53]. The AMPA receptor deactivation rates in both of these cell types are rapid (time constant of deactivation = 1.0 ms at 22 °C) and match the EPSC decay times tightly. The extremely rapid time course of synaptic conductance (< 0.1 ms at physiological temperature), in conjunction with the fast MEMBRANE TIME CONSTANTS, determines that EPSPs generated by activation of these synapses have a short duration. Rapidly decaying EPSPs show little temporal summation, allowing these cells to function as coincidence detectors, which ensure that the timing of action potential firing is constant from one cell to another in these circuits.

This mechanism could theoretically be used to eliminate the location dependence of temporal summation in other neurons. A differential distribution of AMPA receptors could be present in dendrites such that receptor subunits or splice variants showing rapid deactivation (for example, FLOP VARIANTS of GluR2 or 4) would be predominately located at more distal synapses. In cells that use a change in receptor kinetics to reduce location dependence, an increasing density of I_h channels would not be required to normalize temporal summation spatially as long as the membrane time constant is homogeneous throughout the neuron.

TIME INTEGRAL
By accounting for the charge stored in the capacitance of the cells, the time integral of the EPSC relates more to the steady-state voltage attenuation and is therefore much less affected by dendritic filtering.

MAUTHNER CELLS
A bilateral pair of brainstem neurons that receive acoustic information and trigger an escape response characteristic of fish.

MEMBRANE TIME CONSTANTS
These describe the time course of voltage changes in neurons. A small time constant means that the membrane potential can change rapidly.

FLOP VARIANT OF AMPA RECEPTORS
Two splice variants known as flip and flop have been characterized. They differ in their response to glutamate, their distribution and their developmental expression.

© 2000 Macmillan Magazines Ltd

Whereas such a mechanism is possible, it does not seem to be required in CA1 or in neocortical pyramidal neurons. In these cells, the location-dependent changes observed in local EPSP decay kinetics and summation can be completely reproduced by the dendritic injection of EPSC-shaped currents[17,21]. This strongly indicates that a synaptic modification may not be required for the normalization of temporal summation in these cells because an alteration of dendritic membrane properties is able to complete the task. However, this observation does not completely rule out the possibility that distal glutamate receptors desensitize or deactivate differently from proximal receptors. More studies using rapid application techniques are required to compare the kinetics of synaptic glutamate receptors among the different regions of the dendrites, particularly for the more distal synaptic inputs.

Functional impact of location independence

There seem to be several functional benefits associated with eliminating or reducing location-dependent synaptic variability. The normalization of synaptic integration allows a neuron to use Hebbian synaptic mechanisms to store information and it allows this information to be encoded in the temporal firing pattern of a neuronal network. Furthermore, the mechanisms that produce this normalized integration should enrich the repertoire of information processing modes that can be observed in a single neuron.

Hebbian plasticity. When unitary EPSP amplitude and spatio-temporal summation are similar among a group of synapses, each input is equally capable of evoking an action potential. In pyramidal cells, this is the case even at the site of input, because the threshold for action potential initiation increases with distance from the soma at a rate remarkably similar to the increase of the local EPSP amplitude[2,17,33]. As a result, the primary determinant of the weight of a synapse would be its history (whether or not it was associated with the elicitation of an action potential) and not its location in the dendrite. In addition, the generation of action potentials is important for the induction of long-term changes in synaptic strength. Therefore, if all synapses are equally capable of firing the postsynaptic cell, all synapses will have a similar probability of inducing enduring changes of synaptic efficacy such as long-term potentiation or long-term depression. Indeed, when location-dependent synaptic variability is reduced, Hebbian mechanisms can store information efficiently in realistic neuronal networks[54]. If, on the other hand, synapse location were the primary determinant of synaptic efficacy, then changing the weight of a synapse as a function of its association with action potential output would become less efficient for storing information[54]. Therefore, reducing location dependence from synaptic weight can have an extremely positive impact on information storage in neuronal networks.

Links

ENCYCLOPEDIA OF LIFE SCIENCES Dendrites

Population synchrony. The axon of a single CA3 cell can synapse onto the distal dendrites of one CA1 cell and form more proximal synapses on another[45]. Therefore, the spatial innervation pattern of a CA1 pyramidal neuron by any given group of CA3 pyramidal neurons can vary from cell to cell. If the temporal pattern of action potential output from CA1 cells were dependent on the spatial input pattern, it would be difficult to produce synchronous firing from a population of CA1 pyramidal neurons because each cell would receive a spatially different pattern of input. By reducing the location dependence of synaptic integration, a population of CA1 cells that receives widely different spatial patterns of input but similar temporal patterns will fire in synchrony[19]. This synchrony is thought to be an important factor for many functional properties of hippocampal and other neuronal types[55,56].

Integration modes. There are many examples of linear encoding and integration in various sensory systems throughout the brain[57–60]. The reduction of location-dependent synaptic variability and the small amplitude of unitary events could provide a mechanism for the linear summation of activity, where all inputs would contribute equally to the output of the cell. This is, however, only half of the story because ion channels susceptible to modulation (A-type K$^+$-channels, I_h, AMPA receptors and others) are also responsible for the normalization of synaptic weight. In addition, the release of common neuromodulators, such as acetylcholine, noradrenaline and serotonin, at localized regions of the dendrite could lower the firing threshold and alter temporal summation. Furthermore, input patterns are not always highly distributed. So, the appropriate spatio-temporal input pattern in the presence of channel modulation could easily produce a nonlinear form of integration[61–63], perhaps resulting in the local initiation of dendritic spikes[19,33,42]. The same neuron is therefore able to integrate synaptic inputs in an essentially linear or in a highly nonlinear fashion depending on the type of stimuli that it receives. By removing the location dependence of synaptic input, the neuron has extended the range of possible synaptic integration in the same dendritic arbor from a global, uniform type of processing to one that can be regional and highly nonlinear. Furthermore, this feat has been achieved without compromising the ability of the cells to use mechanisms for associative synaptic plasticity and to fire as a synchronous population.

In conclusion, it seems that CA1 pyramidal neurons and probably other nerve cells have normalized the impact of their synaptic inputs through an intricate interplay between morphology, voltage-gated ion channels and the properties of the synapses themselves. Although each of these features shows an amazing complexity, their interplay has, by reducing dendritic distortion, simplified the information carried by tens of thousands of highly distributed dendritic synapses. This simplification can have powerful consequences for the functional capabilities of neurons.

© 2000 Macmillan Magazines Ltd

1 Colbert, C & Johnston, D Axonal action-potential initiation and Na^+ channel densities in the soma and axon initial segment of subicular pyramidal neurons *J. Neurosci.* **16**, 6676–6686 (1996)

2 Stuart, G , Spruston, N , Sakmann, B & Hausser, M Action potential initiation and backpropagation in neurons of the mammalian CNS *Trends Neurosci.* **20**, 125–131 (1997)

3 Rall, W in *Neural Theory and Modeling* (ed Reiss, R F) 5–30 (Stanford Univ Press, Palo Alto, 1964)

4 Rall, W Distinguishing theoretical synaptic potentials computed for different soma-dendritic distributions of synaptic input *J. Neurophysiol* **30**, 1138–1168 (1967)

5 Rall, W Burke, R E , Smith, T G , Nelson, P G & Frank, K Dendritic location of synapses and possible mechanisms for the monosynaptic EPSP in motoneurons. *J. Neurophysiol* **30**, 1169–1193 (1967)

6 Mainen, Z F., Carnevale, N T , Zador, A M , Claiborne, B J & Brown, T H Electrotonic architecture of hippocampal CA1 pyramidal neurons based on three-dimensional reconstruction *J. Neurophysiol.* **76**, 1904–1923 (1996)

7 Rall, W in *Handbook of Physiology. The Nervous System. Cellular Biology of Neurons* Section 1, Vol 1. (eds Kandel E R , Brookhart, J. M & Mountcastle, V B) 39–97 (American Physiological Society, Bethesda, 1977)

8 Jack, J J. B., Noble, D. & Tsein, R. W. in *Electric Current Flow in Excitable Cells* 125–132 (Oxford Univ Press, London, 1975)

9 Segev, I in *The Theoretical Foundation of Dendritic Function* (eds Segev, I , Rinzel, J & Shepard, G. M.) 117–121 (MIT Press, Cambridge, Massachusetts, 1995)

10 Andersen, P., Silfvenius, H., Sundberg, S H. & Sveen, O A comparison of distal and proximal dendrite synapses on CA1 pyramids in guinea pig hippocampal slices *in vitro* *J Physiol (Lond)* **307**, 273–299 (1980)

11 Stricker, C., Field, A. C. & Redman, S J Statistical analysis of amplitude fluctuations in EPSCs evoked in rat CA1 pyramidal neurones *in vitro*. *J. Physiol. (Lond.)* **490**, 419–441 (1996)

12 Sayer, R J., Friedlander, M J & Redman, S J The time course and amplitude of EPSPs evoked at synapses between pairs of CA3/CA1 neurons in the hippocampal slice *J. Neurosci* **10**, 826–836 (1990)

13 Jack, J J , Redman, S. J & Wong, K The components of synaptic potentials evoked in cat spinal motoneurones by impulses in single group Ia afferents. *J. Physiol. (Lond.)* **321**, 65–96 (1981)

14 Inasek, R. & Rèdman, S. J The amplitude, time course and charge of unitary excitatory post-synaptic potentials evoked in spinal motoneurone dendrites. *J. Physiol. (Lond)* **234**, 665–688 (1973)

Provides one of the first and most convincing demonstrations that the location dependence of synaptic amplitude can be reduced by increasing distal synaptic conductance. They have also used sophisticated analysis techniques to reach the same conclusions on the Schaffer collaterals (reference 11).

15 Pettit, D. L & Augustine, G. J. Distribution of functional glutamate and GABA receptors on hippocampal pyramidal cells and interneurons. *J. Neurophysiol.* **84**, 28–38 (2000)

16 Frick, A, Ziegelgansberger, W & Dodt, H U. The subcellular distribution of glutamate receptors in layer V pyramidal neurons. *Soc. Neurosci. Abstrt* **24**, 325 (1998)

17 Magee, J. C & Cook, E P. Synaptic weight is independent of synapse location in CA1 pyramidal neurons *Nature Neurosci* **3**, 895–903 (2000)

18 Magee, J. C Dendritic hyperpolarization-activated currents modify the integrative properties of hippocampal CA1 pyramidal neurons. *J. Neurosci.* **18**, 7613–7624 (1998)

19 Magee, J C. Dendritic I_h normalizes temporal summation in hippocampal CA1 neurons. *Nature Neurosci* **2**, 508–514 (1999)

This paper, followed by reference 17, showed that the location of a synapse has less impact on the amplitude and temporal summation of excitatory postsynaptic potentials (EPSPs) at the soma than expected from theoretical considerations. The location dependence of temporal summation seems to be reduced by the active properties of the dendrites, whereas amplitude seems to be normalized by alterations of synaptic conductance.

20 Andreasen, M & Lambert, J. D. C. Factors determining the efficacy of distal excitatory synapses in rat hippocampal CA1 pyramidal neurones. *J. Physiol. (Lond)* **507**, 421–441 (1998)

21 Williams, S & Stuart, G Site independence of EPSP time course is mediated by dendritic I_h in neocortical neurons *J. Neurophysiol* **83**, 3177–3182 (2000)

22 Cash, S & Yuste, R Linear summation of excitatory inputs by CA1 pyramidal neurons *Neuron* **22**, 383–394 (1999)

Probably the most thorough investigation of spatial summation in a central neuron. The authors compared the summation of iontophoretic potentials at different regions of CA1 neurons and found that summation was mostly linear to slightly sublinear in most dendritic regions. The summation was sensitive to blockade of various ion channels and to the temporal pattern of input, leading the authors to believe that an interaction among voltage-gated ion channels linearizes spatial summation.

23 Rall, W & Rinzel, J Branch input resistance and steady attenuation for input to one branch of a dendritic neuron model *Biophys. J.* **13**, 648–688 (1973)

24 Rinzel, J & Rall, W Transient response in a dendritic neuron model for current injected at one branch *Biophys. J.* **14**, 759–790 (1974)

References 23 and 24 provided new insights into the spread of voltage transients in passive arbours. They describe how the structure of the distal dendritic arbour will increase the local input resistance and provide further paths for current flow. These papers convincingly showed how the local passive properties of the dendrites can produce large-amplitude, kinetically brief local dendritic EPSPs but cannot, however, produce larger or faster EPSPs at the soma.

25 Jaffe, D. B & Carnevale, N T Passive normalization of synaptic integration influenced by dendritic architecture *J. Neurophysiol* **82**, 3268–3285, (1999)

26 Magee, J. C. in *Dendrites* (eds Stuart, G , Spruston, N & Hausser, M) 139–160 (Oxford Univ Press, London, 1999)

27 Wilson, C. J. Dynamic modification of dendritic cable properties and synaptic transmission by voltage-gated potassium channels. *J. Comput Neurosci.* **2**, 91–115 (1995)

28 Stuart, G & Spruston, N Determinants of voltage attenuation in neocortical pyramidal neuron dendrites *J. Neurosci* **18**, 3501–3510 (1998)

29 Cook, E. P & Johnston, D. Voltage-dependent properties of dendrites that eliminate location-dependent variability of synaptic input *J. Neurophysiol* **81**, 535–543 (1999)

30 Gillessen, T. & Alzheimer, C. Amplification of EPSPs by low Ni^{2+}- and amiloride-sensitive Ca^{2+} channels in apical dendrites of rat CA1 pyramidal neurons. *J. Neurophysiol.* **77**, 1639–1643 (1997)

31 Lipowsky, R , Gillessen, T. & Alzheimer, C Dendritic Na^+ channels amplify EPSPs in hippocampal CA1 pyramidal cells. *J. Neurophysiol* **76**, 2181–2191 (1996)

32 Nicoll, A., Larkman, A. & Blakemore, C Modulation of EPSP shape and efficacy by intrinsic membrane conductances in rat neocortical pyramidal neurons *in vitro* *J. Physiol. (Lond)* **468**, 693–705 (1993)

33 Hoffman, D. A., Magee, J. C , Colbert, C. M & Johnston, D K^+ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons *Nature* **387**, 869–875 (1997)

34 Schiller, J , Major, G , Koester, H. J & Schiller, Y NMDA spikes in basal dendrites of cortical pyramidal neurons *Nature* **404**, 285–289 (2000)

35 Urban, N N & Barrionuevo, G Active summation of excitatory postsynaptic potentials in hippocampal CA3 pyramidal neurons *Proc Natl Acad Sci. USA* **95** 11450–11455 (1998)

36 Deisz, R A., Fortin, G & Zieglganberger, W Voltage dependence of excitatory postsynaptic potentials of rat neocortical neurons. *J. Neurophysiol* **65**, 371–382 (1991)

37 Margulis, M & Tang, C Temporal integration can readily switch between sublinear and supralinear summation *J Neurophysiol* **79**, 2809–2813 (1998)

38 Magee, J. C & Johnston, D. A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. *Science* **275**, 209–213 (1997)

39 Larkum, M E , Zhu, J J & Sakmann, B A new cellular mechanism for coupling inputs arriving at different cortical layers *Nature* **398**, 338–341 (1999)

40 Larkum, M. E , Kaiser, K M & Sakmann, B Calcium electrogenesis in distal apical dendrites of layer 5 pyramidal cells at a critical frequency of back-propagating action potentials *Proc Natl Acad. Sci. USA* **96**, 14600–14604 (1999)

41 Bekkers, J M Distribution and activation of voltage-gated potassium channels in cell-attached and outside-out patches from large layer 5 cortical pyramidal neurons of the rat. *J. Physiol. (Lond)* **525**, 611–620 (2000)

42 Golding, N. L & Spruston, N Dendritic sodium spikes are variable triggers of axonal action potentials in hippocampal CA1 pyramidal neurons *Neuron* **21**, 1189–1200 (1998)

43 Larkman, A U., Jack, J J & Stratford, K J Quantal analysis of excitatory synapses in rat hippocampal CA1 *in vitro* during low-frequency depression *J. Physiol (Lond)* **505**, 457–471 (1997)

44 Lim, R , Alvarez, F J & Walmsley, B Quantal size is correlated with receptor cluster area at glycinergic synapses in the rat brainstem *J. Physiol. (Lond.)* **516** 505–512 (1999)

45 Sorra, K E & Harris, K M Occurrence and three-dimensional structure of multiple synapses between individual radiatum axons and their target pyramidal cells in hippocampal area CA1 *J. Neurosci.* **13**, 3736–3748 (1993)

46 Nusser, Z. *et al.* Cell type and pathway dependence of synaptic AMPA receptor number and variability in the hippocampus *Neuron* **21**, 545–559 (1998)

47 Takumi, Y., Ramirez-León, V, Laake P , Rinvik, E & Ottersen, O P Different modes of expression of AMPA and NMDA receptors in hippocampal synapse *Nature Neurosci* **2**, 618–624 (1999)

48 Nusser, Z , Cull-Candy, S. & Farrant, M Differences in synaptic GABA(A) receptor number underlie variation in GABA mini amplitude *Neuron* **19**, 697–709, (1997)

49 Korn, H., Bausela, F., Charpier S. & Faber, D S Synaptic noise and multiquantal release at dendritic synapses *J. Neurophysiol* **70**, 1249–1254 (1993)

Shows that the amplitude distal synaptic currents on goldfish Mauthner neurons are increased by the release of more than one single quantum of transmitter. The authors discuss evidence for an increased number of synaptic boutons with multiple release sites as the underlying mechanism for the normalization of the amplitude.

50 Spruston, N., Jonas, P. & Sakmann, B Dendritic glutamate receptor channels in rat hippocampal CA3 and CA1 pyramidal neurons. *J. Physiol. (Lond)* **482**, 325–352 (1995)

51 Otmakhova, N A , Otmakhov, N & Lisman, J Is NMDA/AMPA ratio higher in perforant path–CA1 input? *Soc Neurosci Abstr* **26**, 351 (2000)

52 Geiger, J. R. P , Lubke, J , Roth, A., Frotscher, M. & Jonas, P Submillisecond AMPA receptor-mediated signaling at a principal neuron–interneuron synapse *Neuron* **18**, 1009–1023 (1997)

Using an impressive combination of techniques, this paper shows how the kinetics of glutamate receptors can influence temporal integration of excitatory postsynaptic potentials (EPSPs). A specific receptor subunit composition produces rapid deactivation of α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptors and a greatly reduced temporal summation.

53 Trussell, L. O. Synaptic mechanisms for coding timing in auditory neurons *Annu. Rev. Physiol* **61**, 477–496 (1999)

54 Cook, E P. & Johnston, D Active dendrites reduce location-dependent variability of synaptic input trains *J. Neurophysiol.* **78**, 2116–2128 (1997)

55 Buzsaki, G & Chrobak, J J. Temporal structure in spatially organized neuronal ensembles a role for interneuronal networks. *Curr. Opin. Neurobiol.* **5**, 504–510 (1995)

56 O'Keefe, J & Recce, M. L Phase relationship between hippocampal place units and the EEG theta rhythm *Hippocampus* **3**, 317–330 (1993)

57 Powers, R. K & Binder, M D Summation of effective synaptic currents and firing rate modulation in cat spinal motoneurons. *J. Neurophysiol* **83**, 483–500, (2000)

58 Ferster, D. Linearity of synaptic interactions in the assembly of receptive fields in the cat visual cortex *Curr. Opin. Neurobiol* **4**, 563–568 (1994)

59 Deangelis, G. C , Ohzawa I & Freeman, R. D Spatiotemporal organization of simple-cell receptive fields in the cats striate cortex II Linearity of temporal and spatial summation. *J. Neurophysiol* **69**, 1118–1135 (1993)

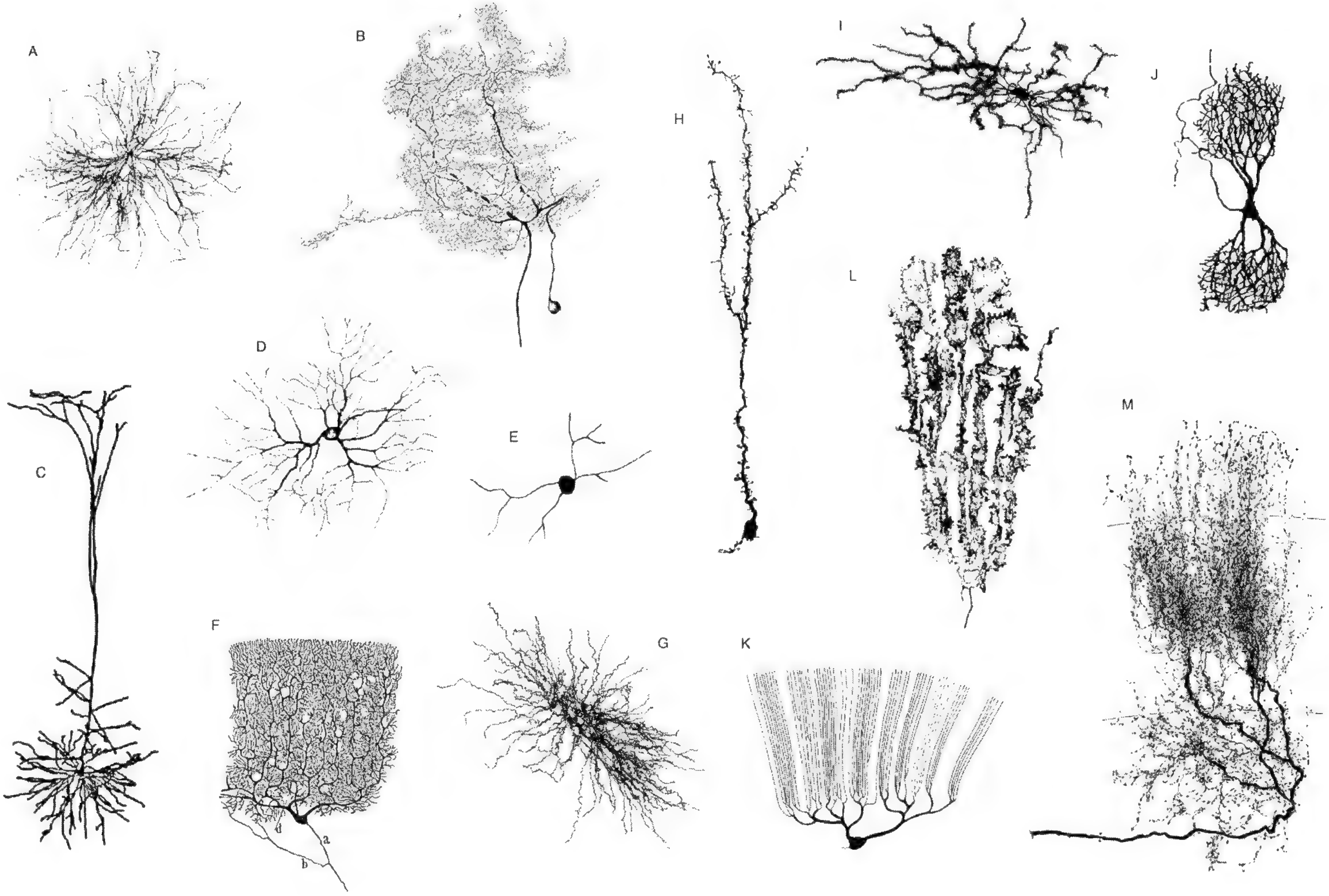
60 Steinmetz, M. A , Motter, B. C , Duffy, C. J & Mountcastle, V B. Functional properties of parietal visual neurons. radial organization of directionalities within the visual field *J. Neurosci* **7**, 177–191 (1987)

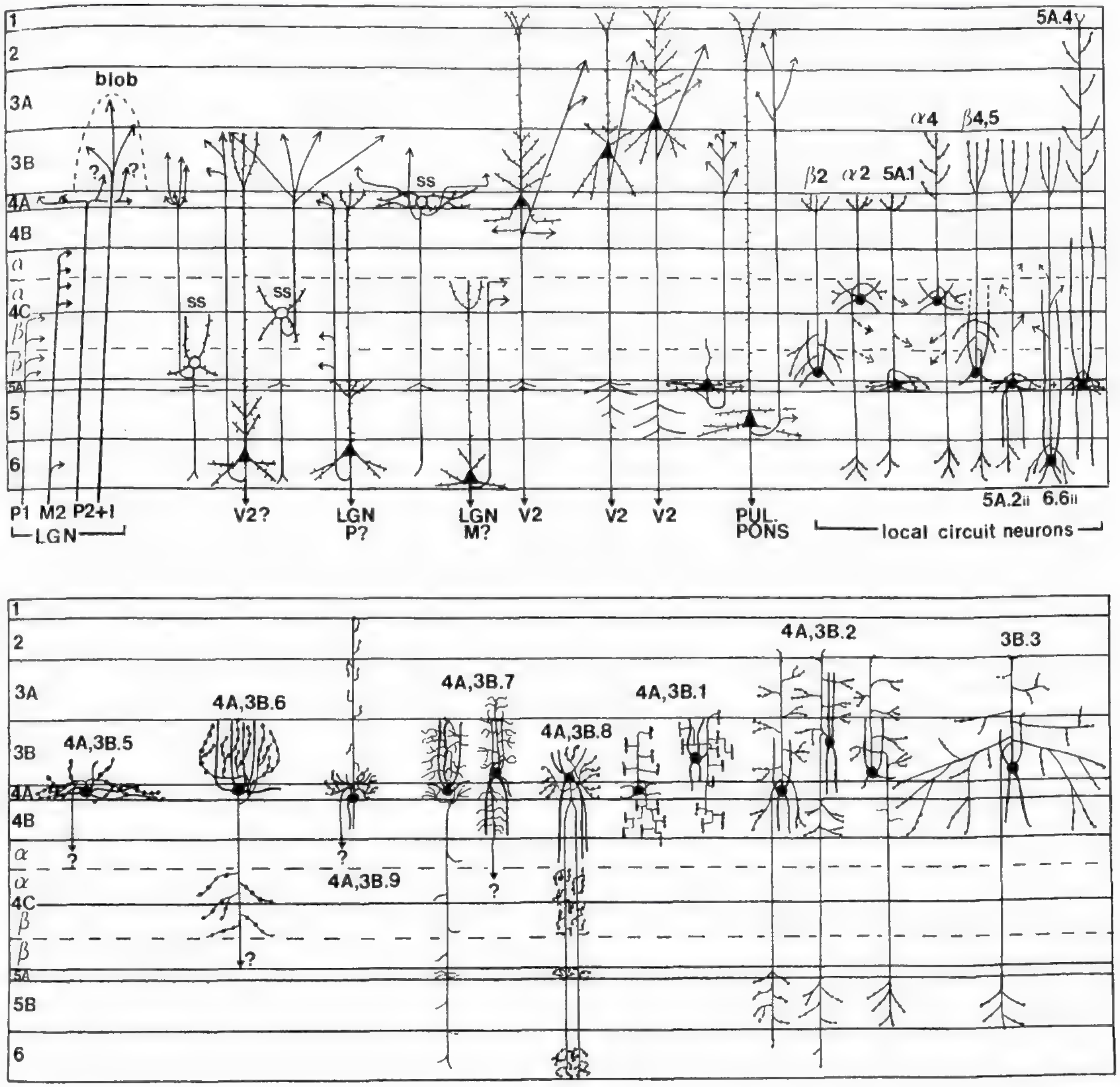
61 Johnston, D , Hoffman, D., Colbert, C & Magee, J. C Modulation of backpropagating action potentials in dendrites of hippocampal pyramidal neurons *Curr. Opin. Neurobiol.* **9**, 288–292 (1999)

62 Mel, B. Synaptic integration in an excitable dendritic tree *J. Neurophysiol.* **70**, 1086–1101(1993)

63 Archie, K A. & Mel, B W A model for intradendritic computation of binocular disparity *Nature Neurosci.* **3**, 54–63, 2000

© 2000 Macmillan Magazines Ltd


A
B
C
D
E
F
G
H
I
J
K
L
M
a
b
d

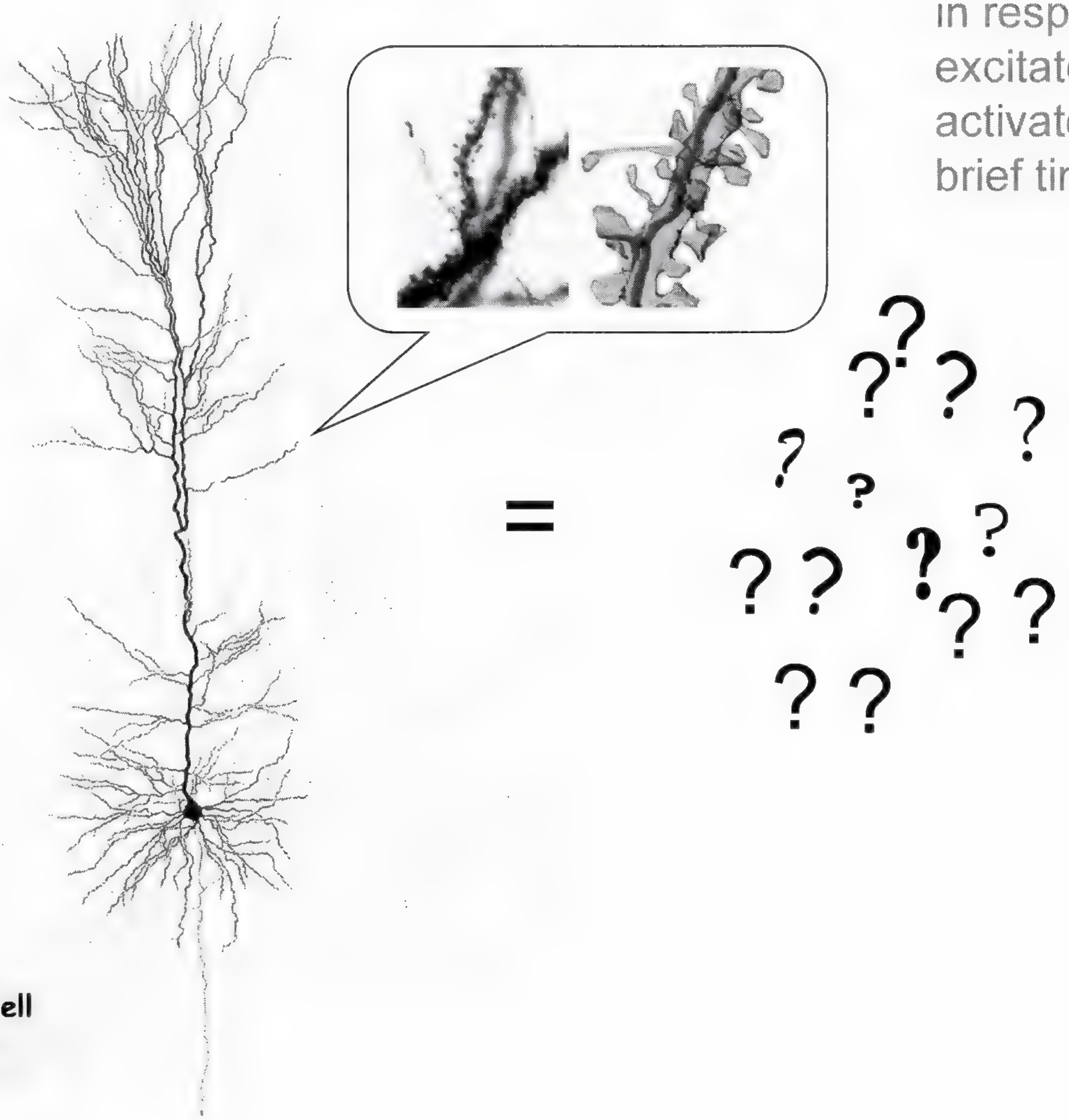

1
2
3A
3B
4A
4B
α
α
4C
β
β
5A
5
6
5B
blob
ss
5A.4
α4
β4,5
β2
α2
5A.1
P1 M2 P2+I
LGN
V2?
LGN P?
LGN M?
V2
V2
V2
PUL. PONS
5A.2ii
6.6ii
local circuit neurons
4A,3B.5
4A,3B.6
4A,3B.9
4A,3B.7
4A,3B.8
4A,3B.1
4A,3B.2
3B.3
?

Lund & Yoshioko (1991)

single, thin dendrite of a

What does a pyramidal neuron do?

in response to
excitatory Inputs
activated during a
brief time window

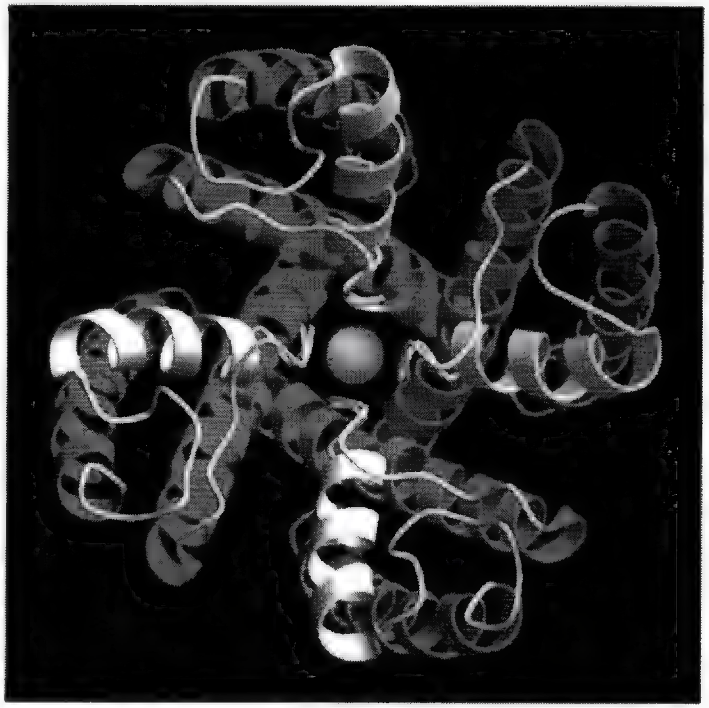


Layer 5 pyramidal cell
(Lubke et al. 1996)

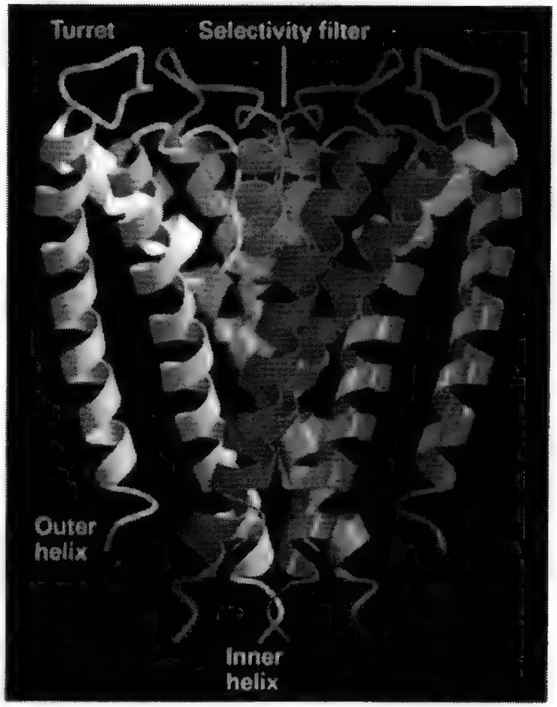
Electrical (circuit) preliminaries

Ion channels

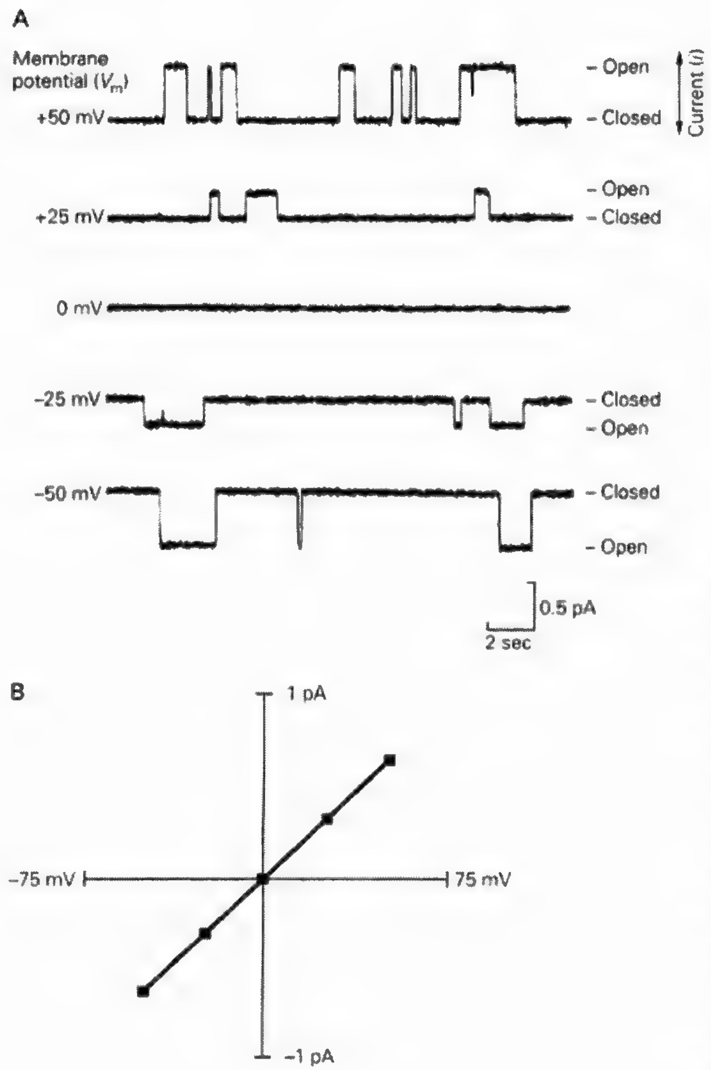
top view



side view


Turret
Selectivity filter
Outer helix
Inner helix

Ion channel as linear resistor



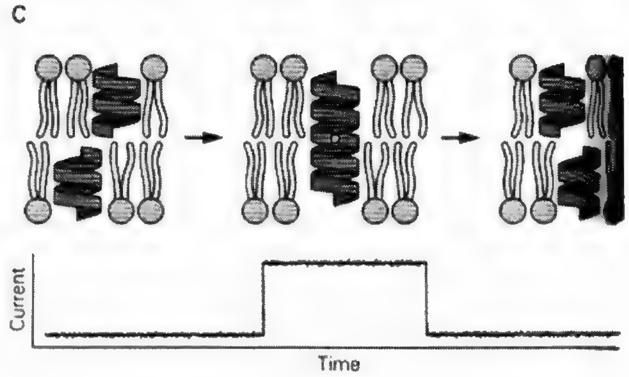


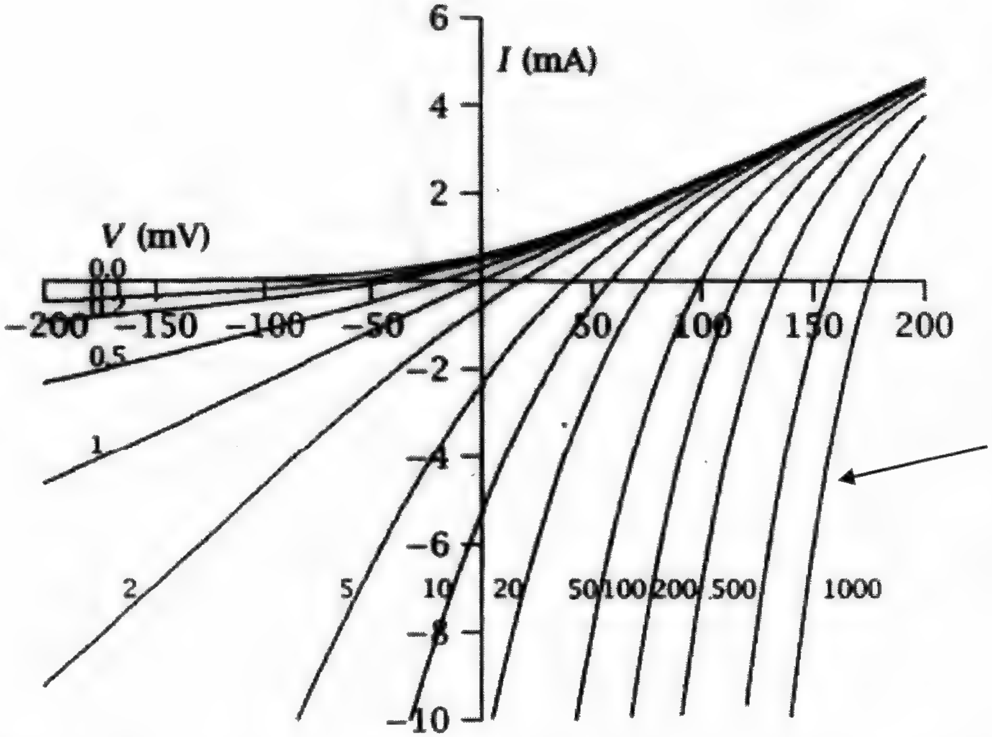
Figure 6-2 Characteristics of the current in a single ion **channel.** The data presented here were obtained from a cha nel formed by the addition of gramicidin A molecules to the s lution bathing an artificial lipid bilayer.

A. The channel opens and closes in an all-or-none fashion, resulting in brief current pulses through the membrane. If the electrical potential (V_m) across the membrane is varied, the cu rent through the channel (i) changes proportionally. V_m is mea sured in millivolts (mV); i is measured in picoamperes (pA).

B. A plot of the current through the channel versus the potential difference across the membrane reveals that the current linearly related to the voltage; in other words, the channel behaves as an electrical resistor that follows Ohm's law ($i = V/R$ or $i = \gamma \times V$). (Data courtesy of Olaf Anderson and Lyndon Providence.)

C. Proposed structure of the gramicidin A channel. A functio channel is formed by end-to-end dimerization of two gramic peptides. (From Sawyer et al. 1989.)

Not really linear, but we can pretend…



Acts like higher conductance when flowing from high [*C*] to low [*C*]

Figure 2.5 Current-voltage relations given by equation 2.7.17 (GHK current equation) for various values of $[C]_{out}/[C]_{in}$ (indicated by small numbers near each curve).

The “battery”

TABLE 3. FREE IONIC CONCENTRATIONS AND EQUILIBRIUM POTENTIALS FOR MAMMALIAN SKELETAL MUSCLE

Ion	Extracellular concentration (m*M*)	Intracellular concentration (m*M*)	$\frac{[\mathrm{Ion}]_o}{[\mathrm{Ion}]_i}$	Equilibrium potential[a] (mV)	neurons (mV)
Na^+	145	12	12	+67	+55
K^+	4	155	0.026	−98	−75
Ca^{2+}	1.5	$<10^{-7}$ M	>15,000	> +128	+115
Cl^-	123	4.2[b]	30[b]	−90[b]	−60

[a] Calculated from Equation 1-11 at 37°C.
[b] Calculated assuming a −90-mV resting potential for the muscle membrane and that Cl^- ions are at equilibrium at rest.

[Hille, 1984]

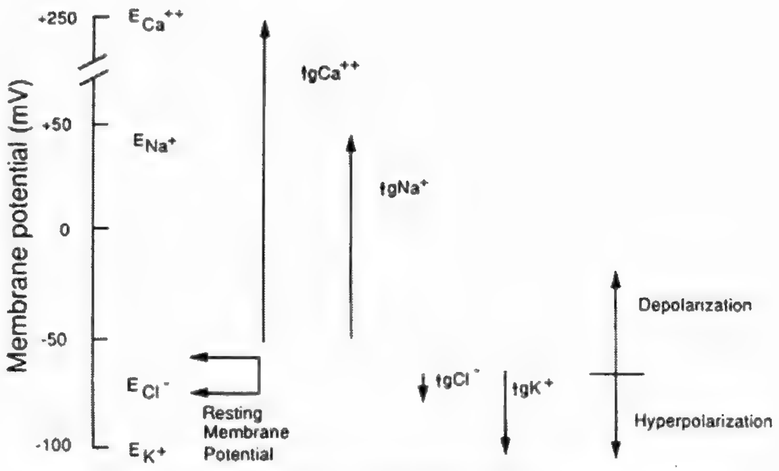
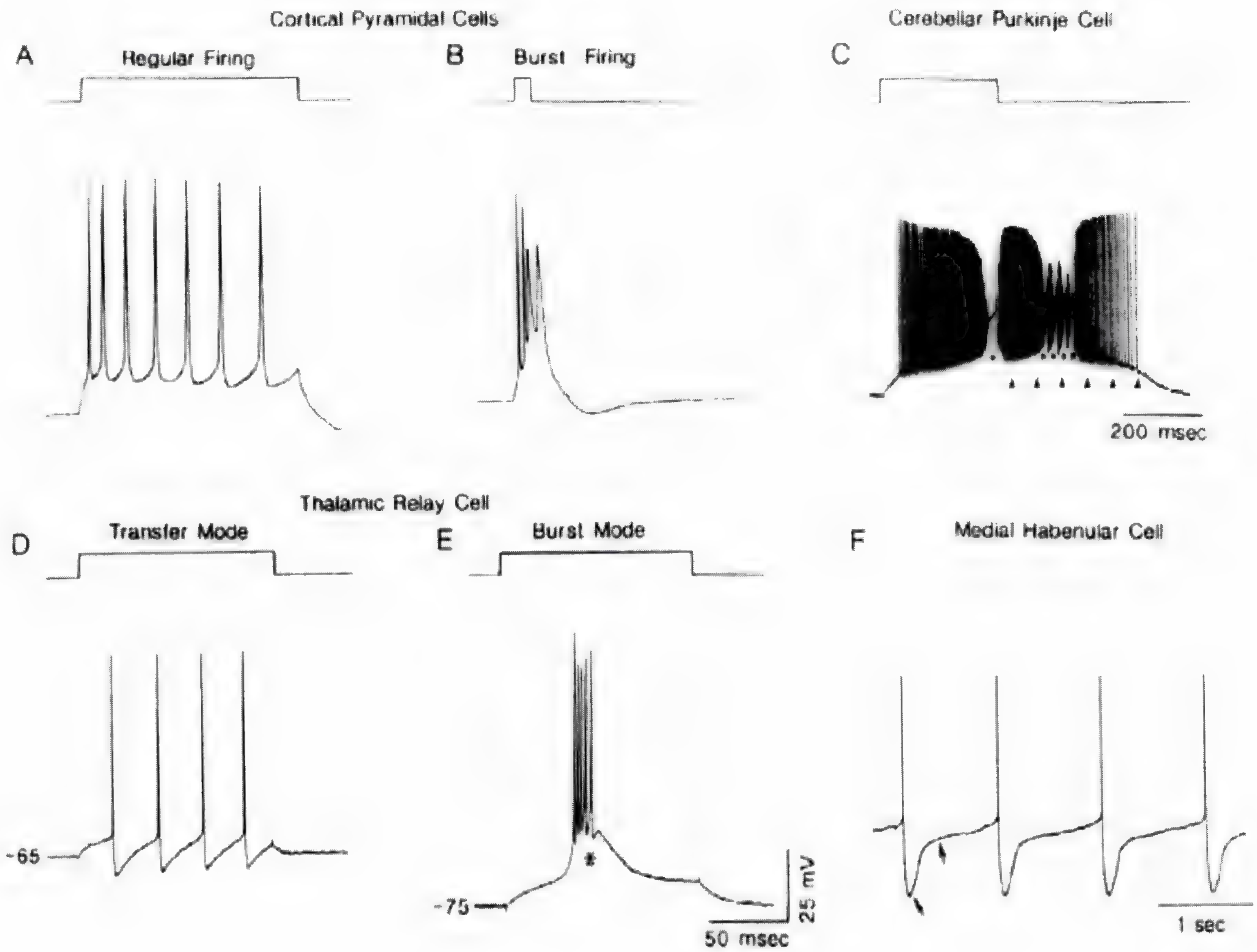

Membrane potential (mV)
+250
+50
0
-50
-100
$E_{Ca^{++}}$
E_{Na^+}
E_{Cl^-}
E_{K^+}
Resting Membrane Potential
↑gCa++
↑gNa+
↑gCl−
↑gK+
Depolarization
Hyperpolarization

FIG. 2.2. Effect of increasing membrane conductance (denoted as *g*) to Ca^{2+}, Na^+, Cl^-, or K^+. Increases in gCa^{2+} or gNa^+ bring the membrane potential toward more positive values (*depolarization*), whereas increases in gCl^- or gK^+ bring it toward more negative values (*hyperpolarization*).

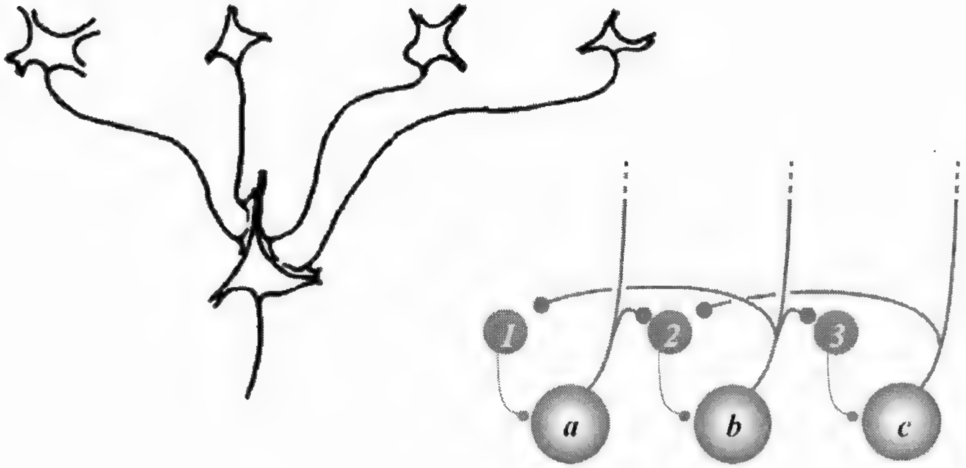
[McCormick, 1990]

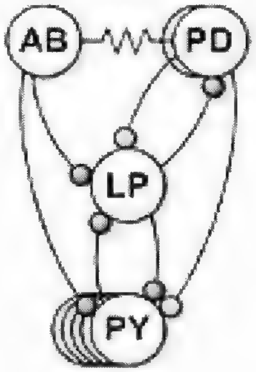
Adding active channels -> I&F neuron

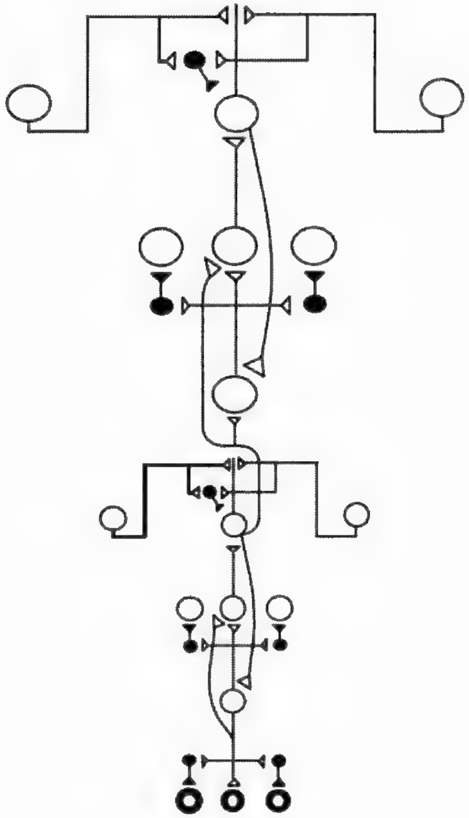
Spikes of the world

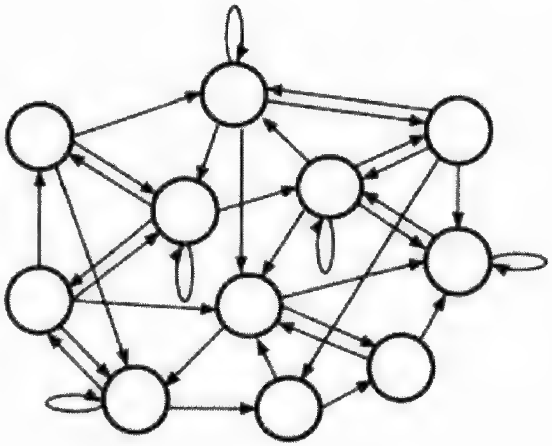


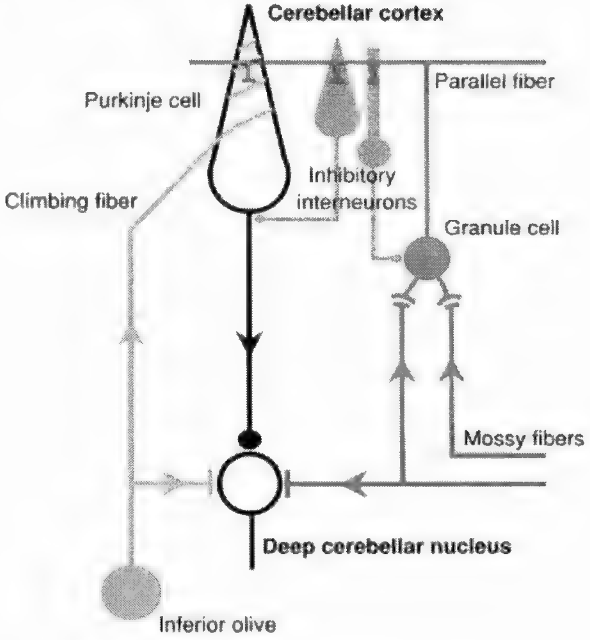
“Point” neurons: the historical favorite





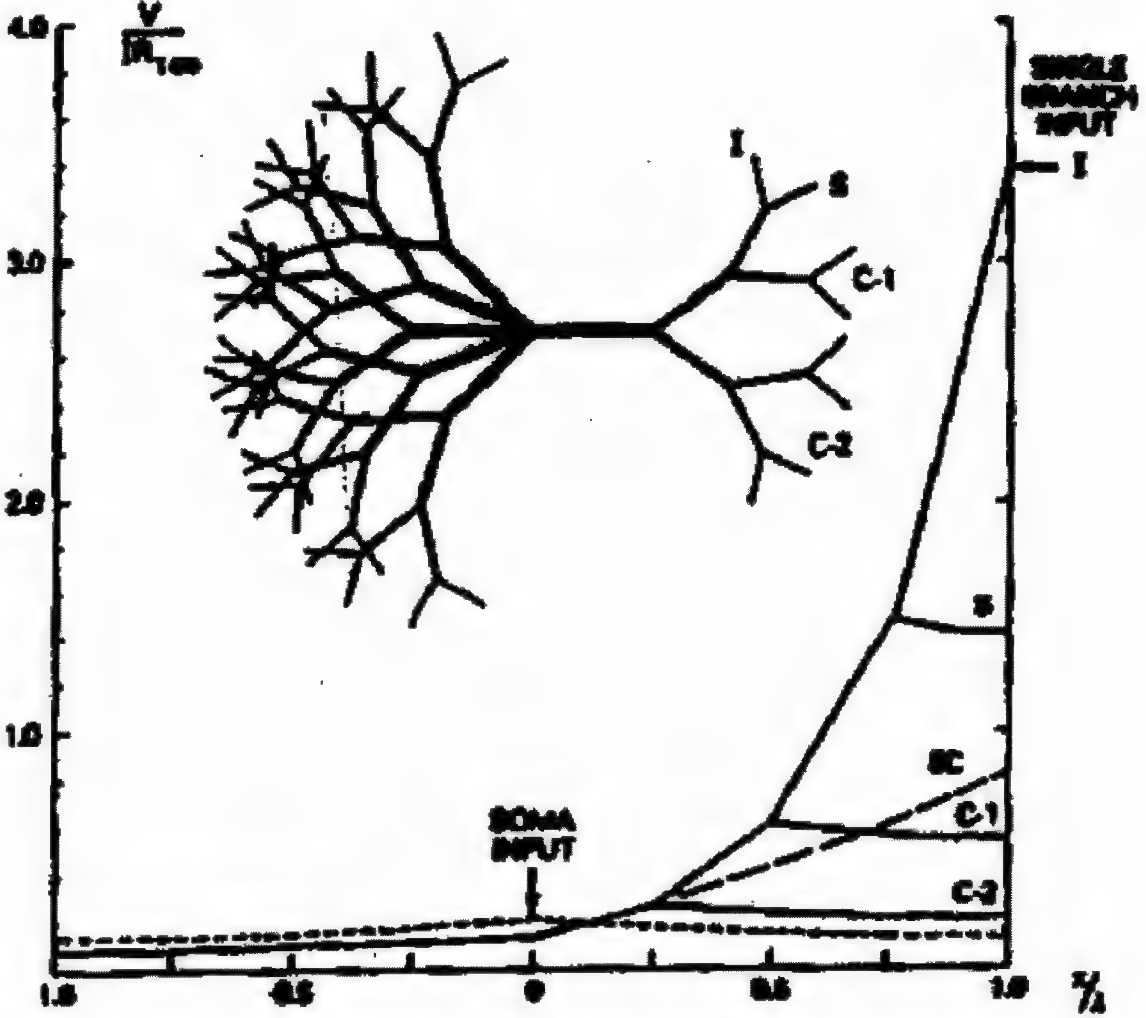




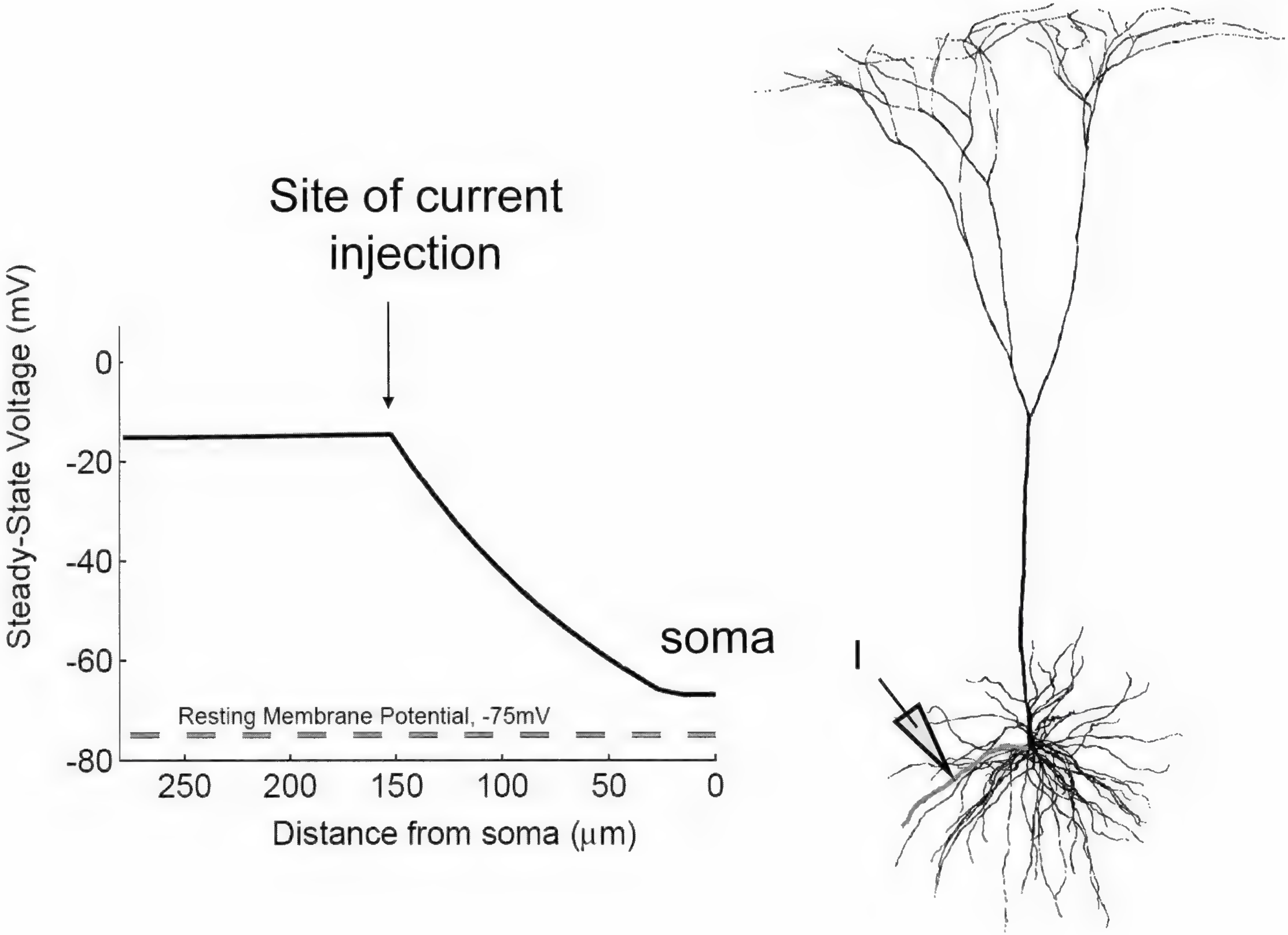

Cerebellar cortex
Purkinje cell
Parallel fiber
Inhibitory interneurons
Climbing fiber
Granule cell
Mossy fibers
Deep cerebellar nucleus
Inferior olive

Some cable basics

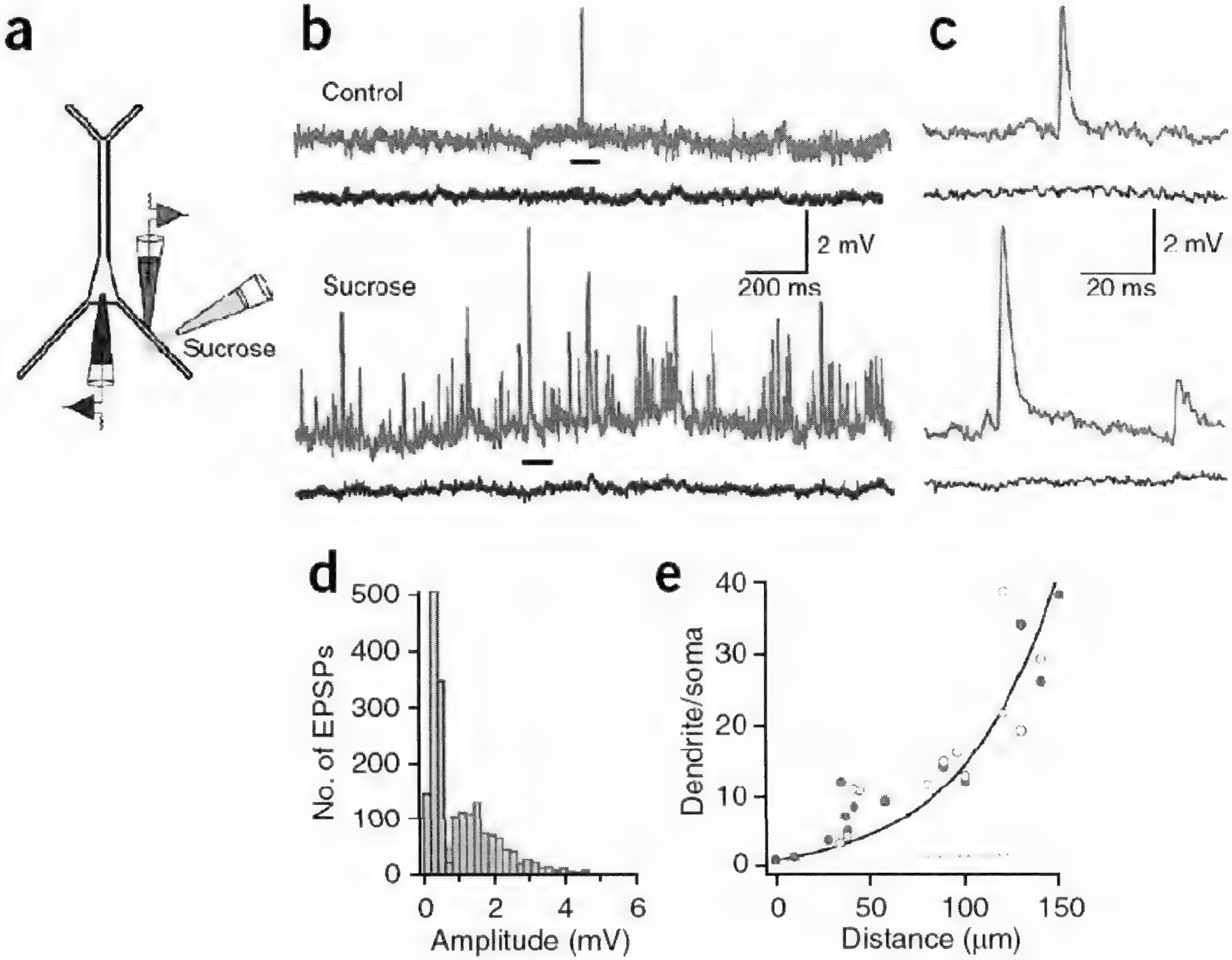
Asymmetric voltage attenuation



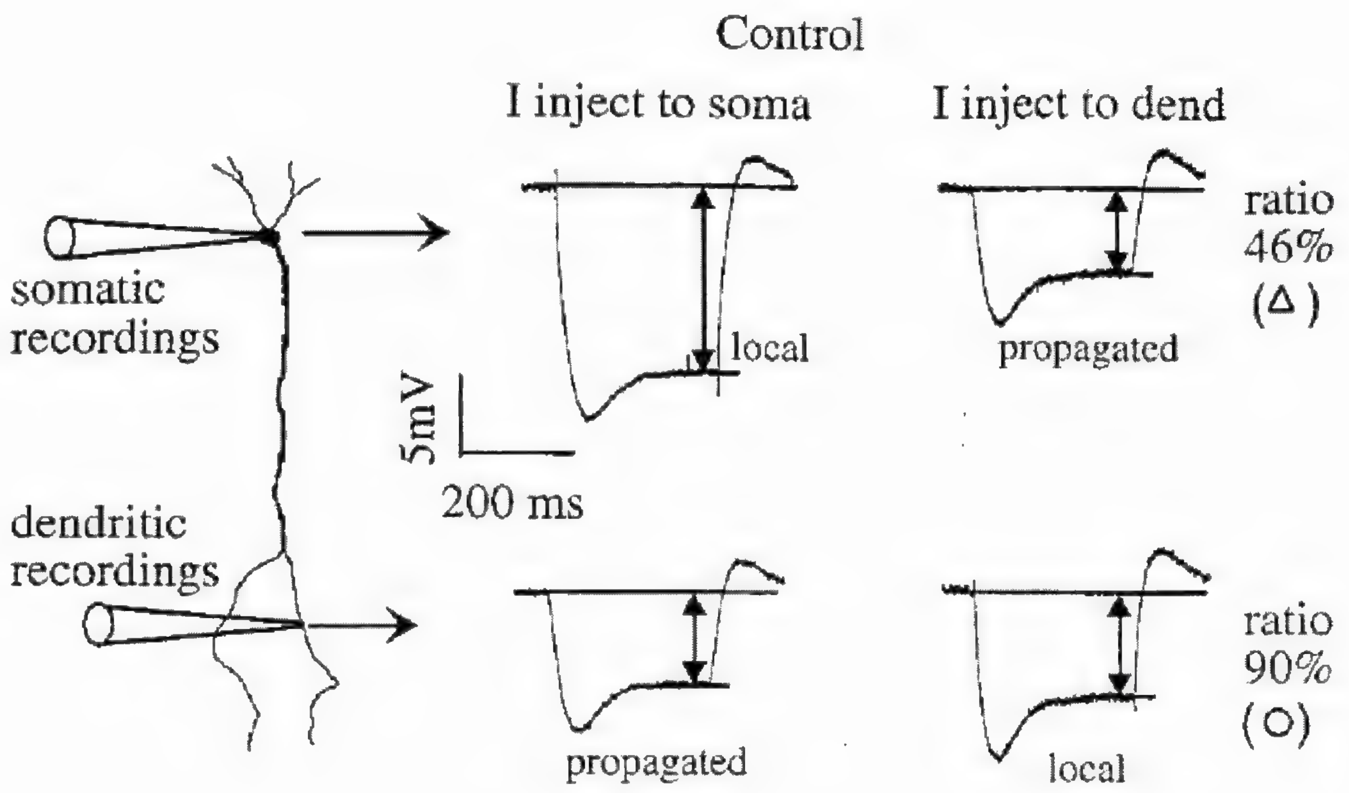
Rall & Rinzel (1973)


Site of current injection
Steady-State Voltage (mV)
0
-20
-40
-60
-80
soma
Resting Membrane Potential, -75mV
250
200
150
100
50
0
Distance from soma (μm)
I

Nevian, Larkum, Polsky & Schiller (2006)



Caution: attenuation isn't always so asymmetric…



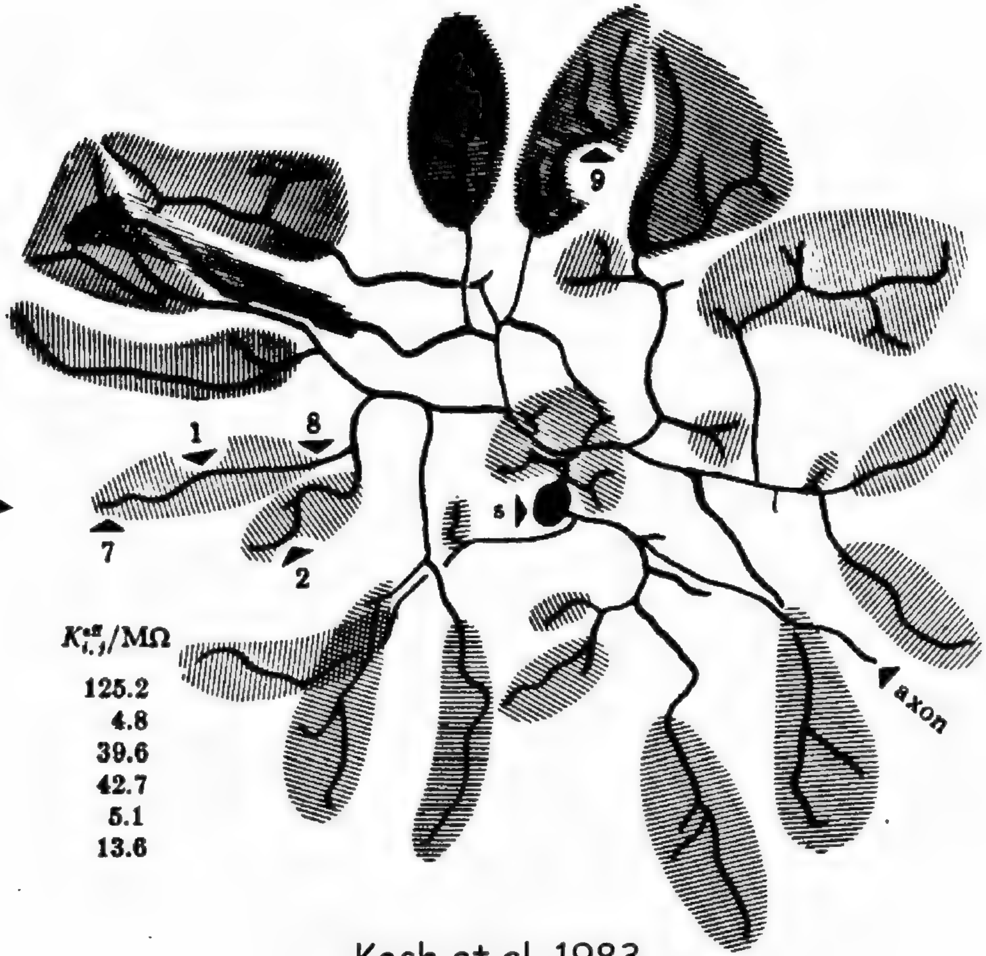
Magee (1998)

First tricky issue

Dendritic "subunits"

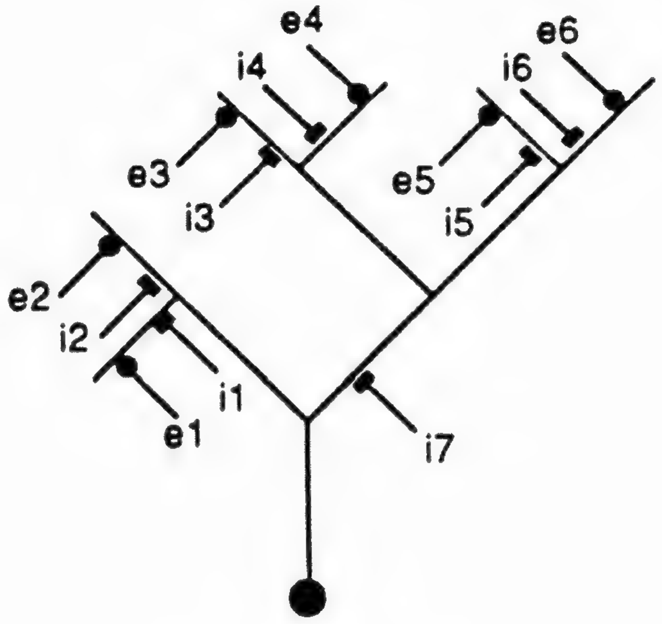
Conclusion: dendrites are compartmentalized into “subunits”

Early computer modeling study showed that dendrites can provide a good deal of electrical isolation for synaptic inputs giving rise to quasi-independent dendritic "subunits"


9
1
8
7
2
s
axon
$K^{ss}_{i,j}$/MΩ
125.2
4.8
39.6
42.7
5.1
13.6

Koch et al. 1983

Dendritic subunits + nonlinearity = world domination


e4
i4
e6
i6
e3
i3
e5
i5
e2
i2
i1
e1
i7

Koch et al. (1983)

Dendritic spikes

NMDA spikes

letters to nature

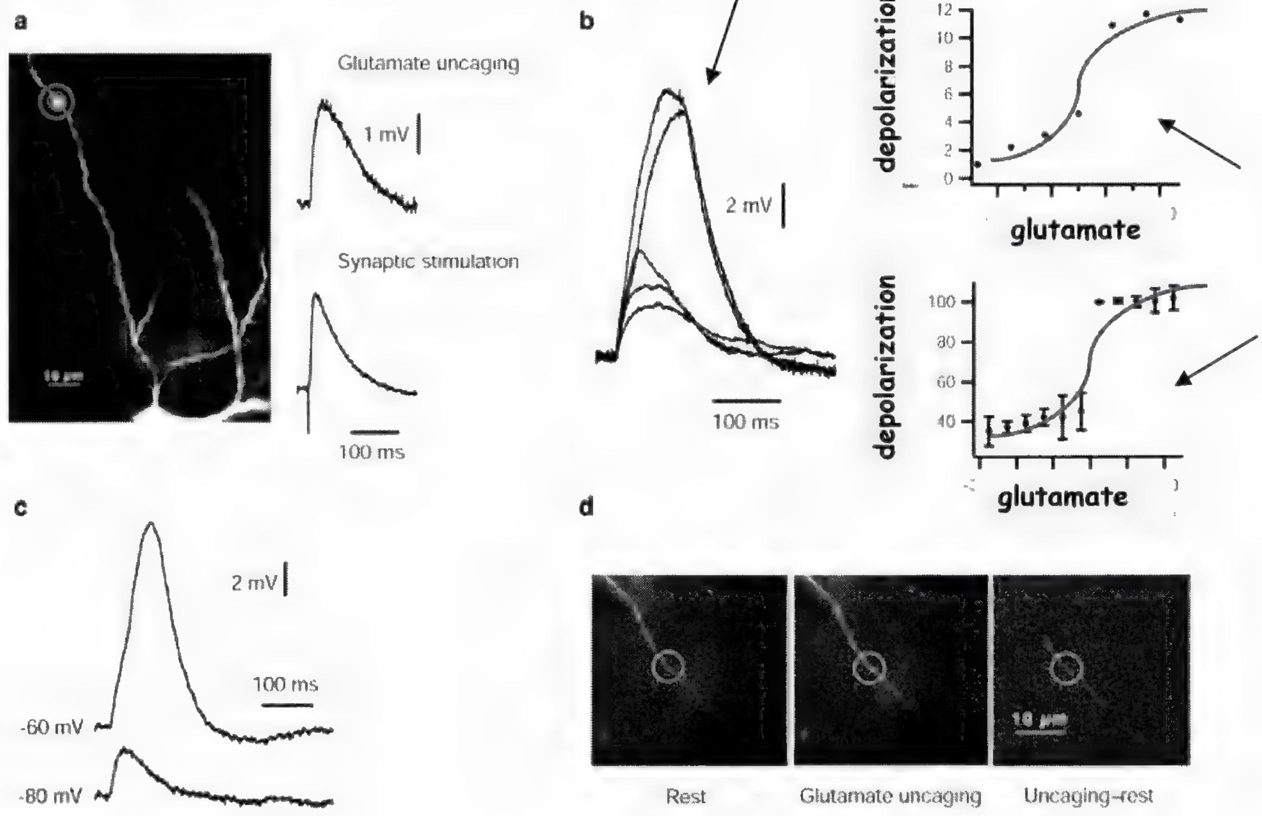
NATURE VOL 404 16 MARCH 2000 www.nature.com

NMDA spikes in basal dendrites of cortical pyramidal neurons

Jackie Schiller*, Guy Major≤≥, Helmut J. Koester≥ß & Yitzhak Schiller≥k

Laser uncaging of glutamate activates a local cluster of synapses

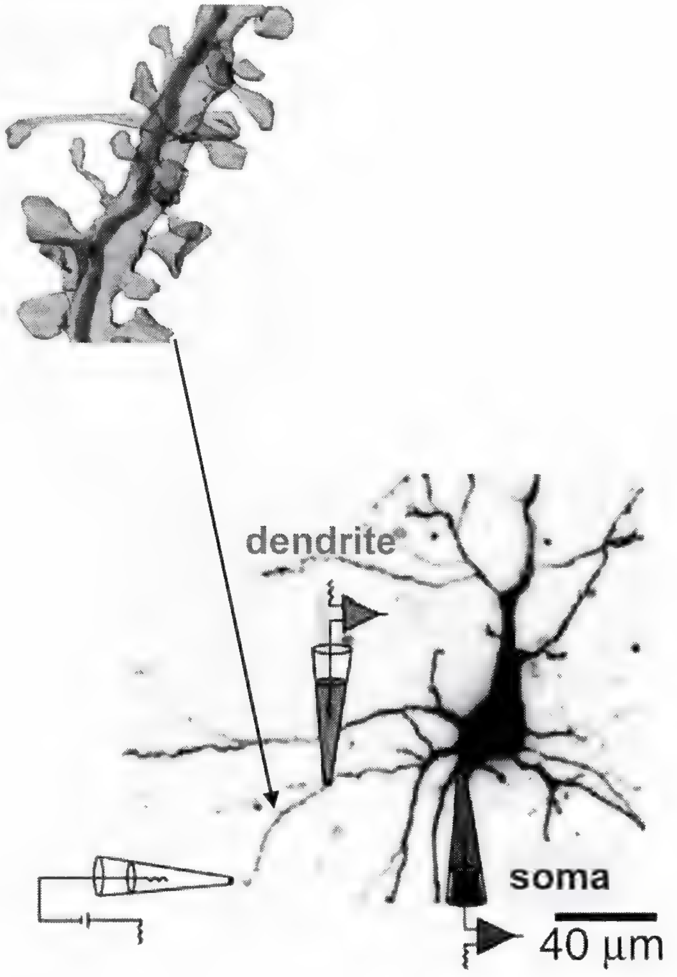
Big jump in somatic EPSP seen when certain threshold is reached



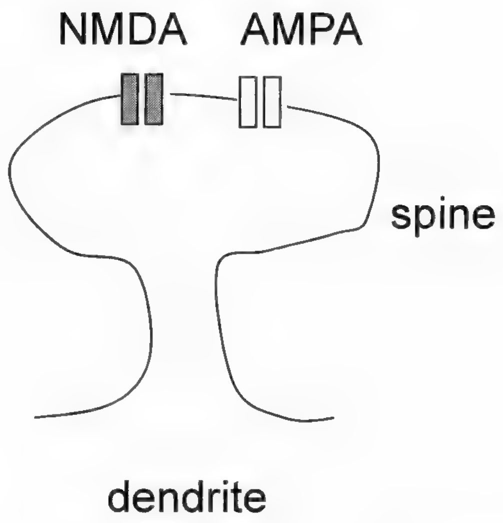
Sigmoidal nonlinearity

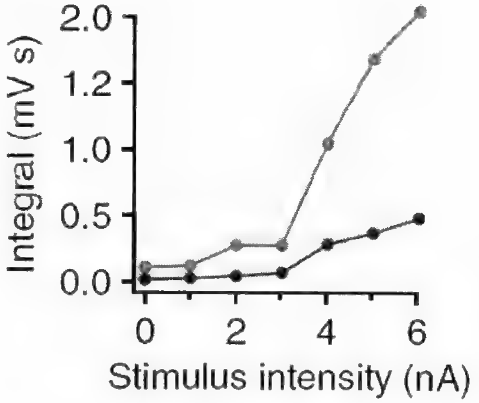
NMDA (and Na+) spikes: a closer look

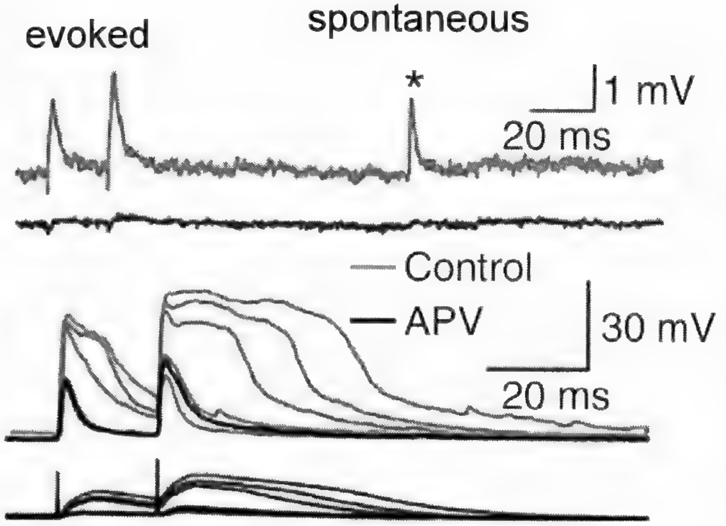
Kristen Harris


dendrite
soma
40 µm

Nevian, Larkum, Polsky & Schiller (2006)

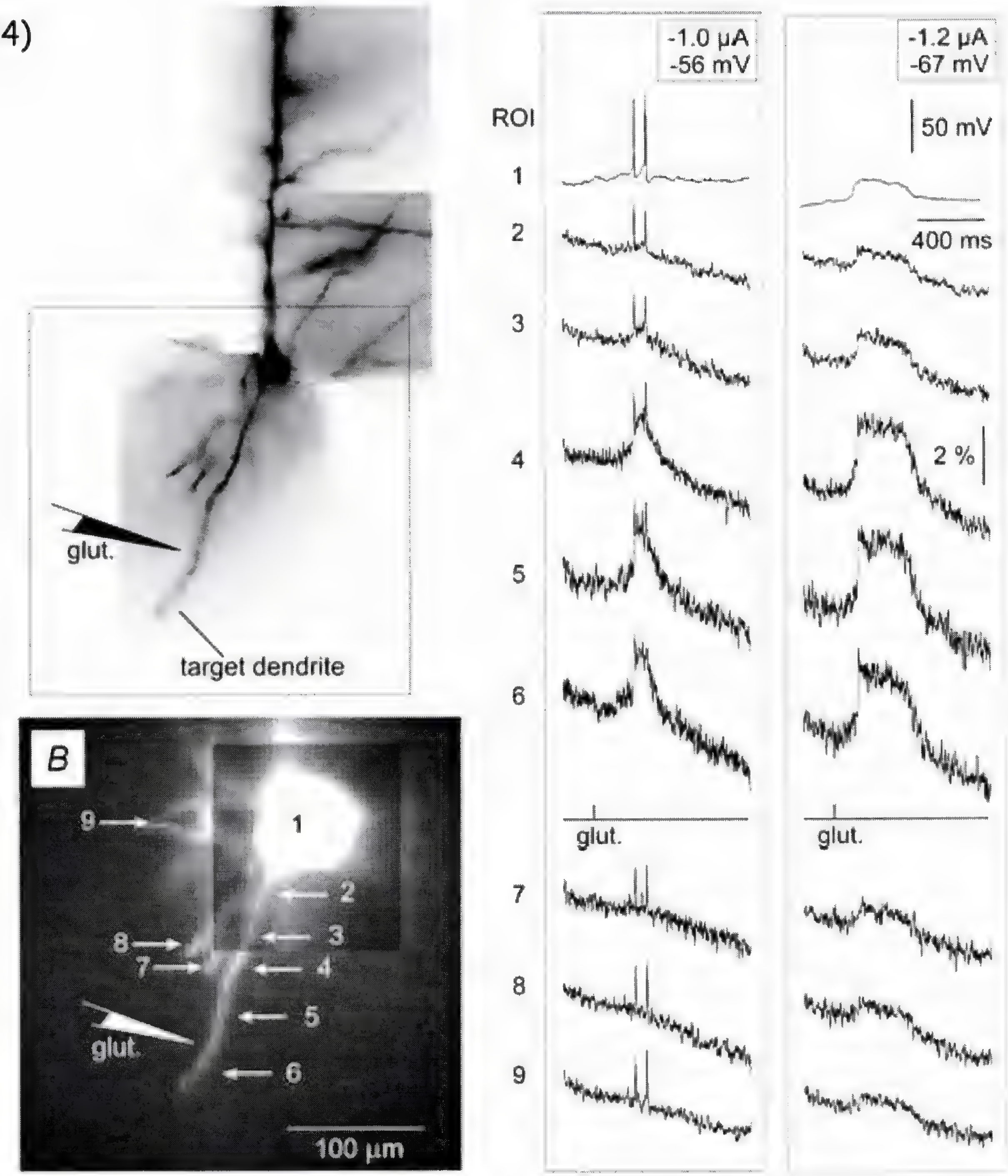

NMDA
AMPA
spine
dendrite





Milojkovic et al. (2004)

Using voltage-
sensitive
dyes to record local
regenerative spikes
in thin branches

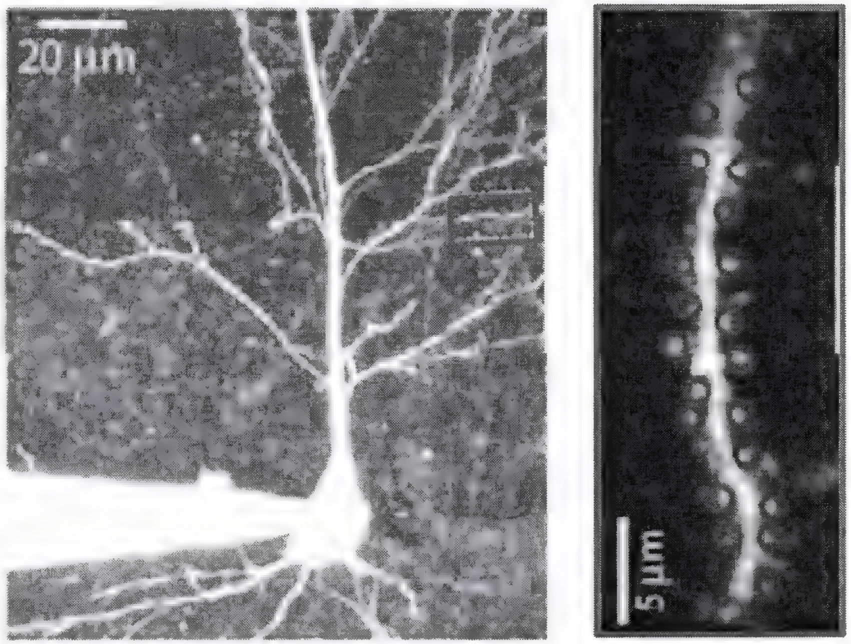


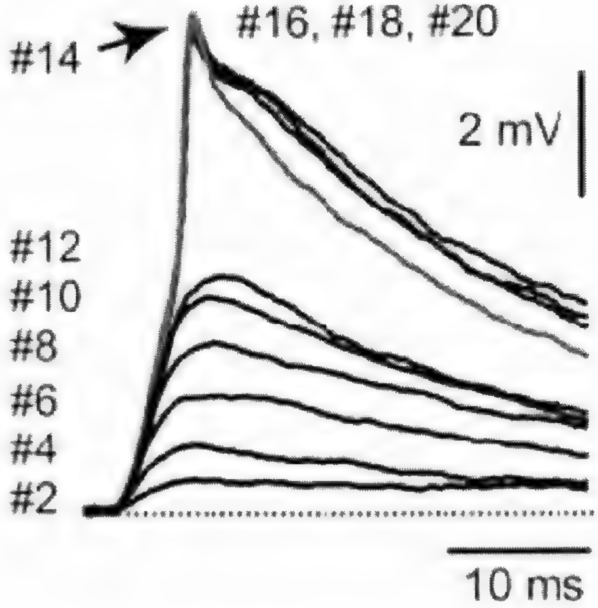
spikes
are well
confined
within
single
branches

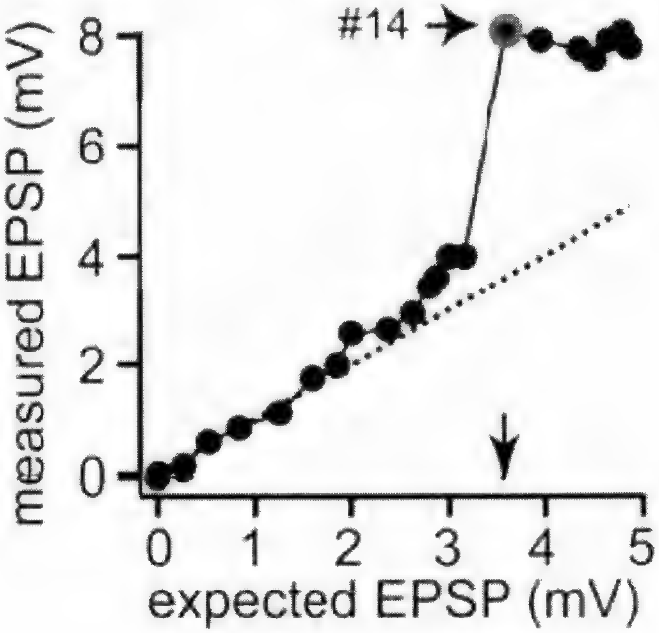
neighbor
branches

Losonczy & Magee (2006)

CA1



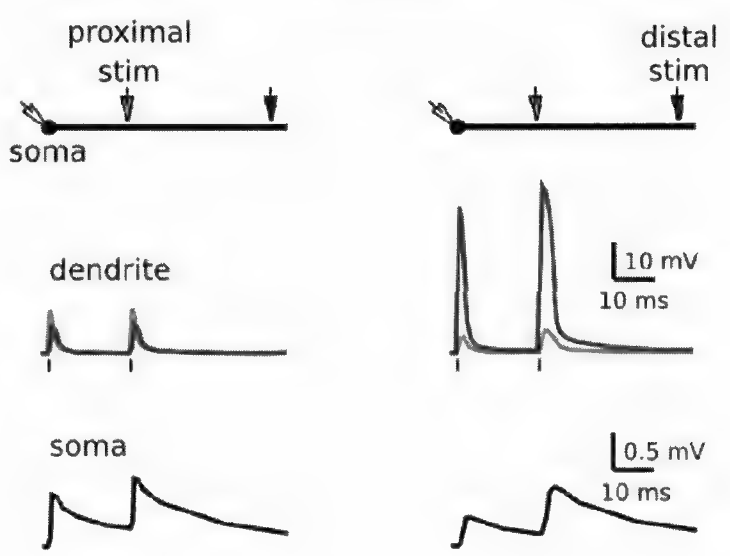




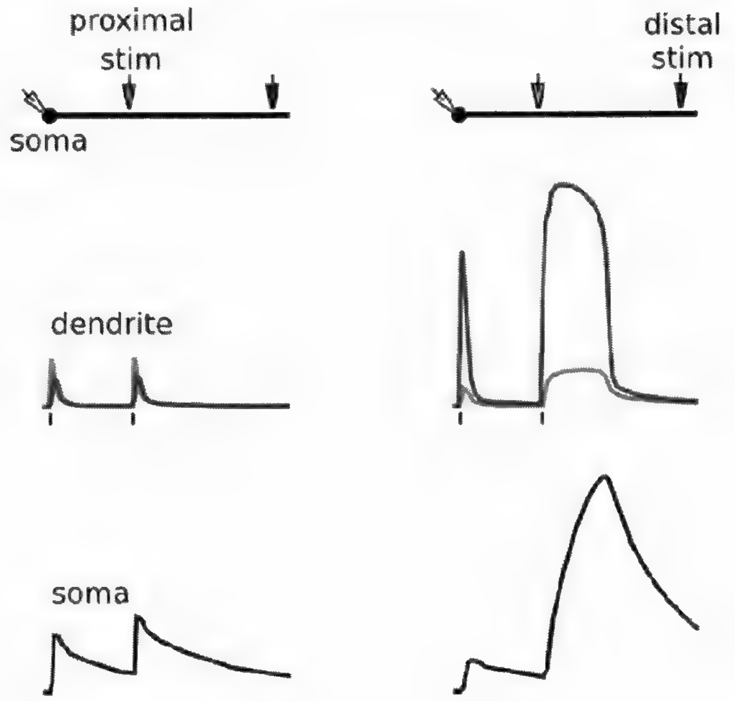
. . . That the local spikes appear to initiate near the middle of the input pattern, even though the input is spread across nearly the entire branch, further supports the idea that all branch input is similarly effective in initiating local spikes. Together, these findings demonstrate that oblique branches can be single integrative compartments that summate without regard to the spatial distribution of the input within the branch.

With dendritic spikes, things get a bit more complicated

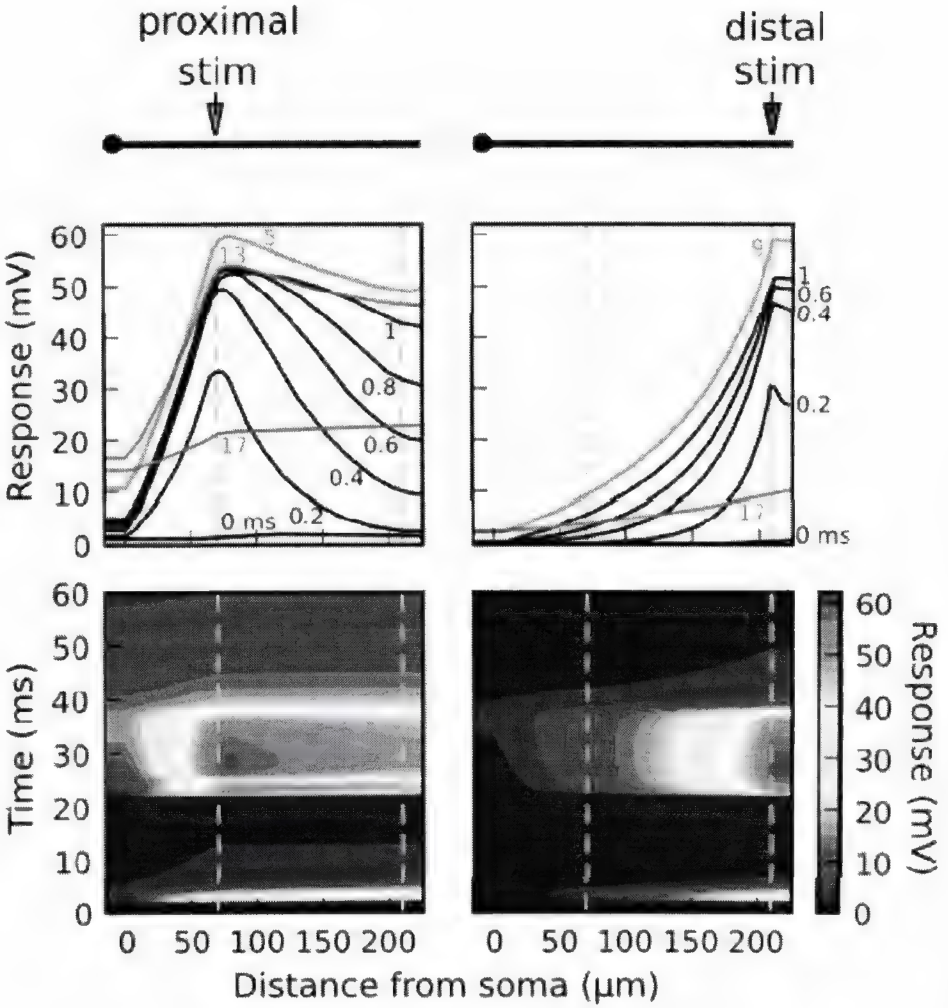
7 synapses



8 synapses

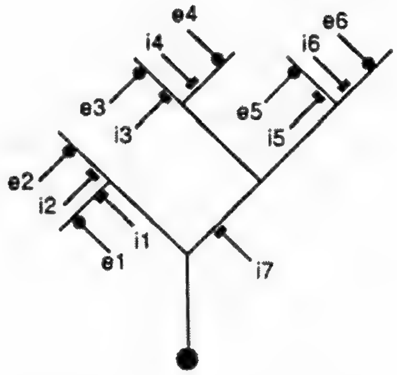


“space plot”

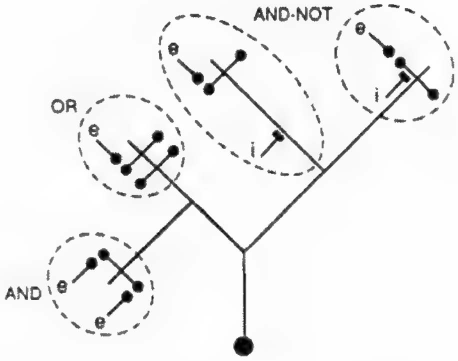


Toward a "network" model of the single neuron

Early models of dendritic function

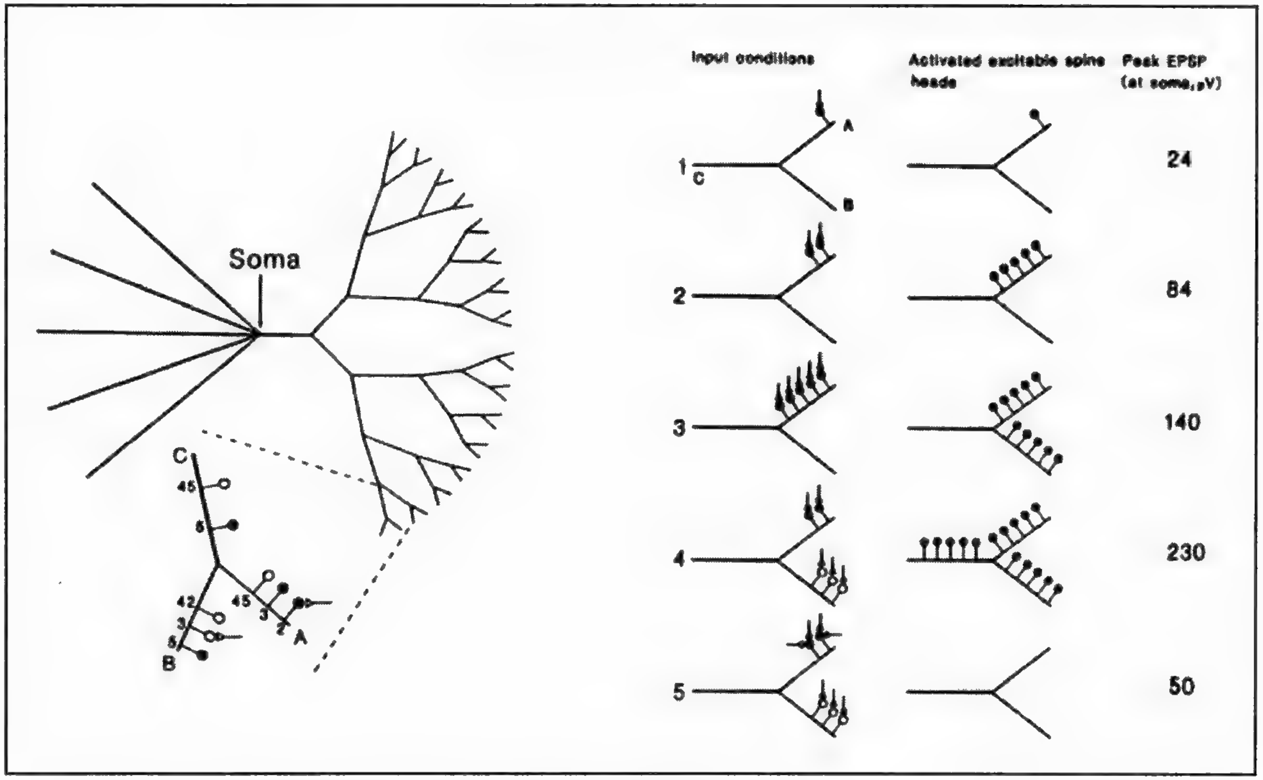

e4
e6
i4
i6
e3
e5
i3
i5
e2
i2
i1
e1
i7

Koch et al. (1983)


AND-NOT
e
e
OR
e
i
i
AND
e
e

Shepherd & Koch (1985)

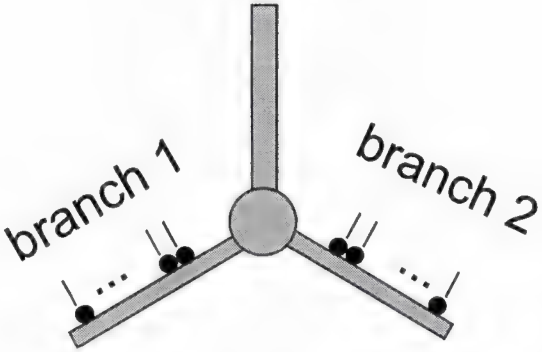
First simulation of active dendrites driven by complex synaptic stimuli


Soma
C
45
8
42
3
5
B
45
3
2
A
Input conditions
Activated excitable spine heads
Peak EPSP (at soma, μV)
1
C
A
B
24
2
84
3
140
4
230
5
50

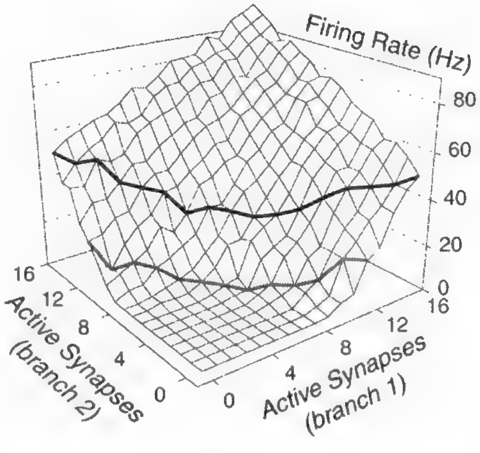
Rall & Segev (1987)

The sum-of-squares model

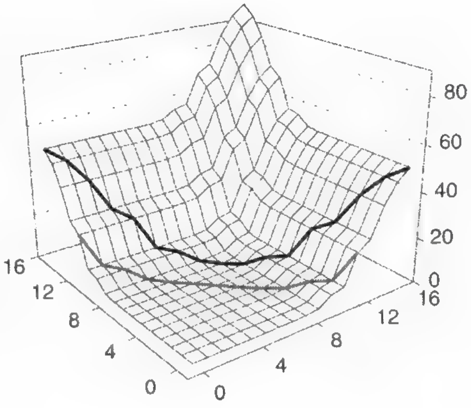
Kevin Archie



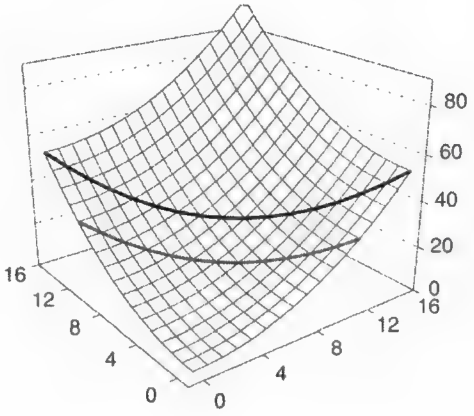
100 Hz input trains
AMPA/NMDA synapses
Na+/K+ channels



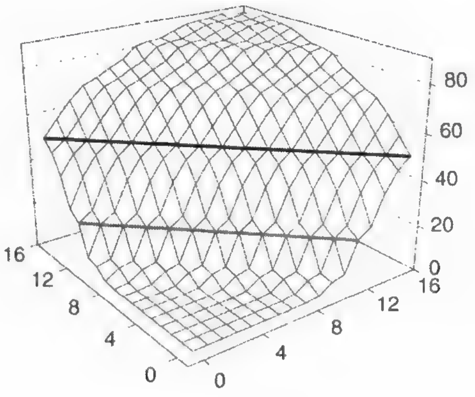
Two Branch Experiment



Hypothetical:
Perfectly isolated subunits



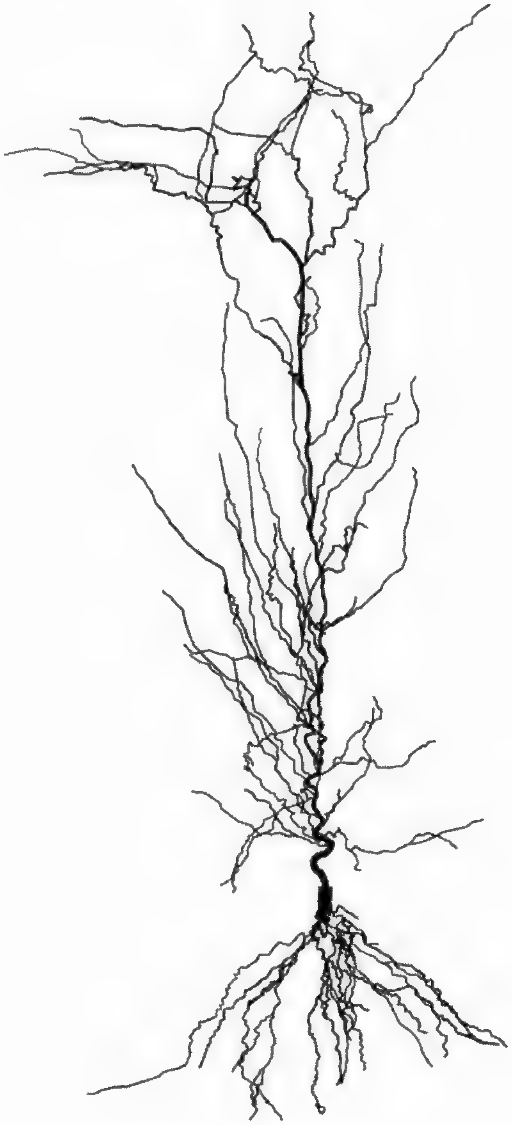
Hypothetical:
Half-squaring subunits



Hypothetical:
Linear subunits

A CA1 pyramidal cell model

Yiota Poirazi



Morphology

CA1 pyramidal cell

183 compartments

Channels

I_{Na}

I_{Kdr}

I_{A}

I_{m}

I_{h}

I_{CaT}

$I_{CaRsoma}$

$I_{CaRdend}$

$I_{CaLsoma}$

$I_{CaLdend}$

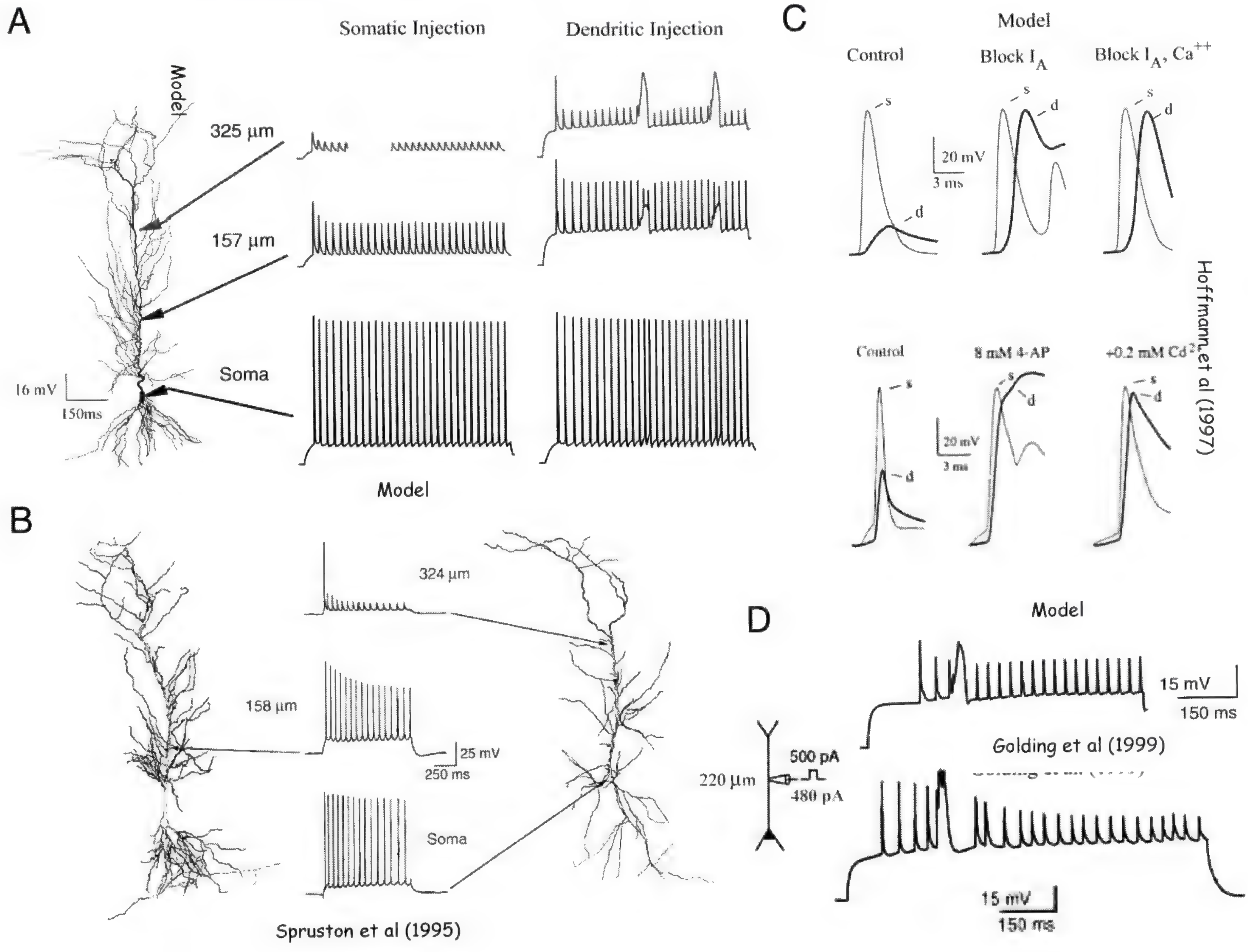
I_{sAHP}

I_{mAHP}

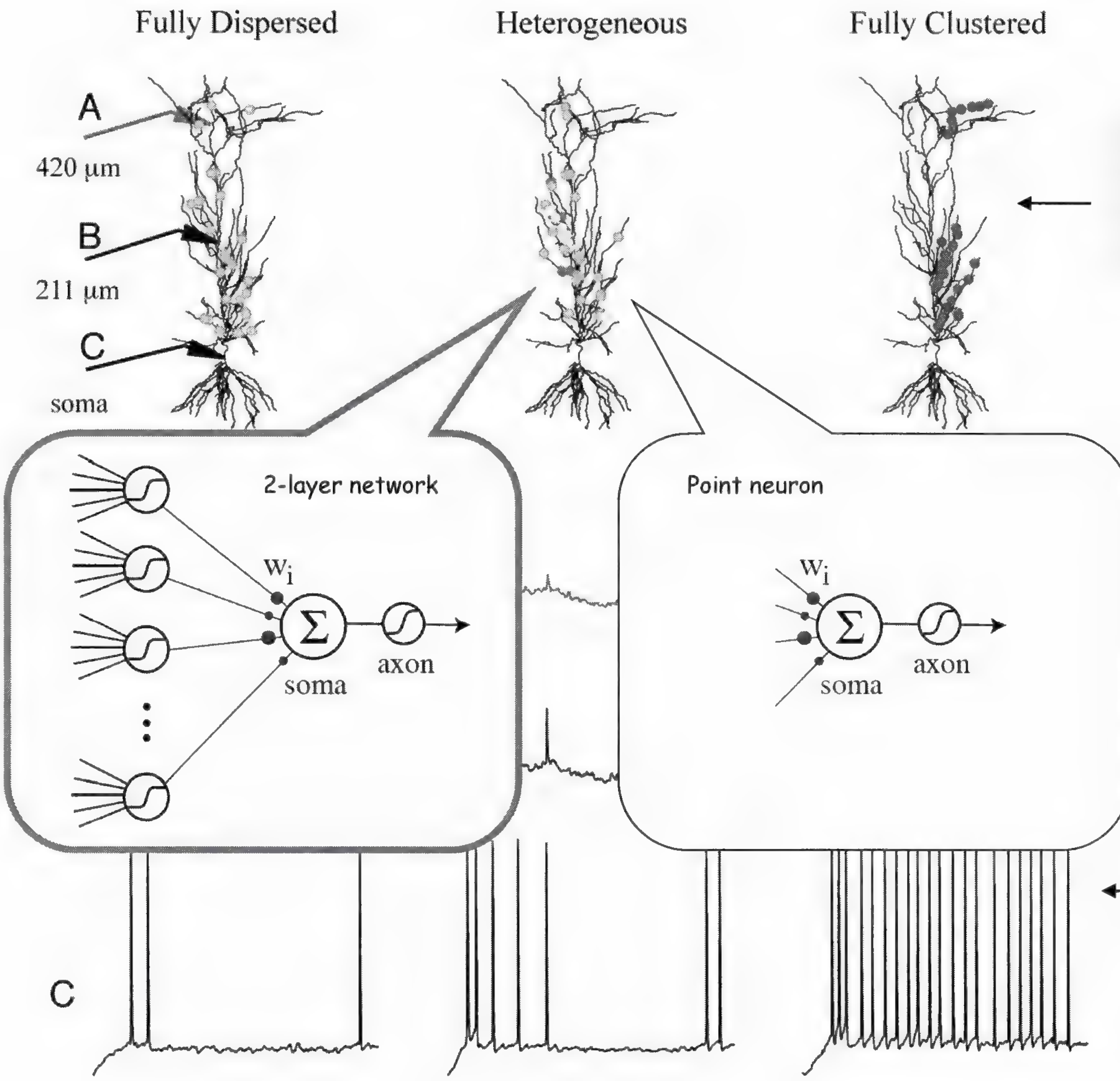
Synapses

NMDA, AMPA,

$GABA_A$, $GABA_B$


A
Somatic Injection
Dendritic Injection
Model
325 µm
157 µm
Soma
16 mV
150ms
Model
B
324 µm
158 µm
25 mV
250 ms
Soma
Spruston et al (1995)
C
Model
Control
Block I_A
Block I_A, Ca^{++}
s
d
20 mV
3 ms
Control
8 mM 4-AP
+0.2 mM Cd^{2+}
20 mV
3 ms
Hoffmann et al (1997)
D
Model
15 mV
150 ms
Golding et al (1999)
220 µm
500 pA
480 pA
15 mV
150 ms

Predicting firing rates

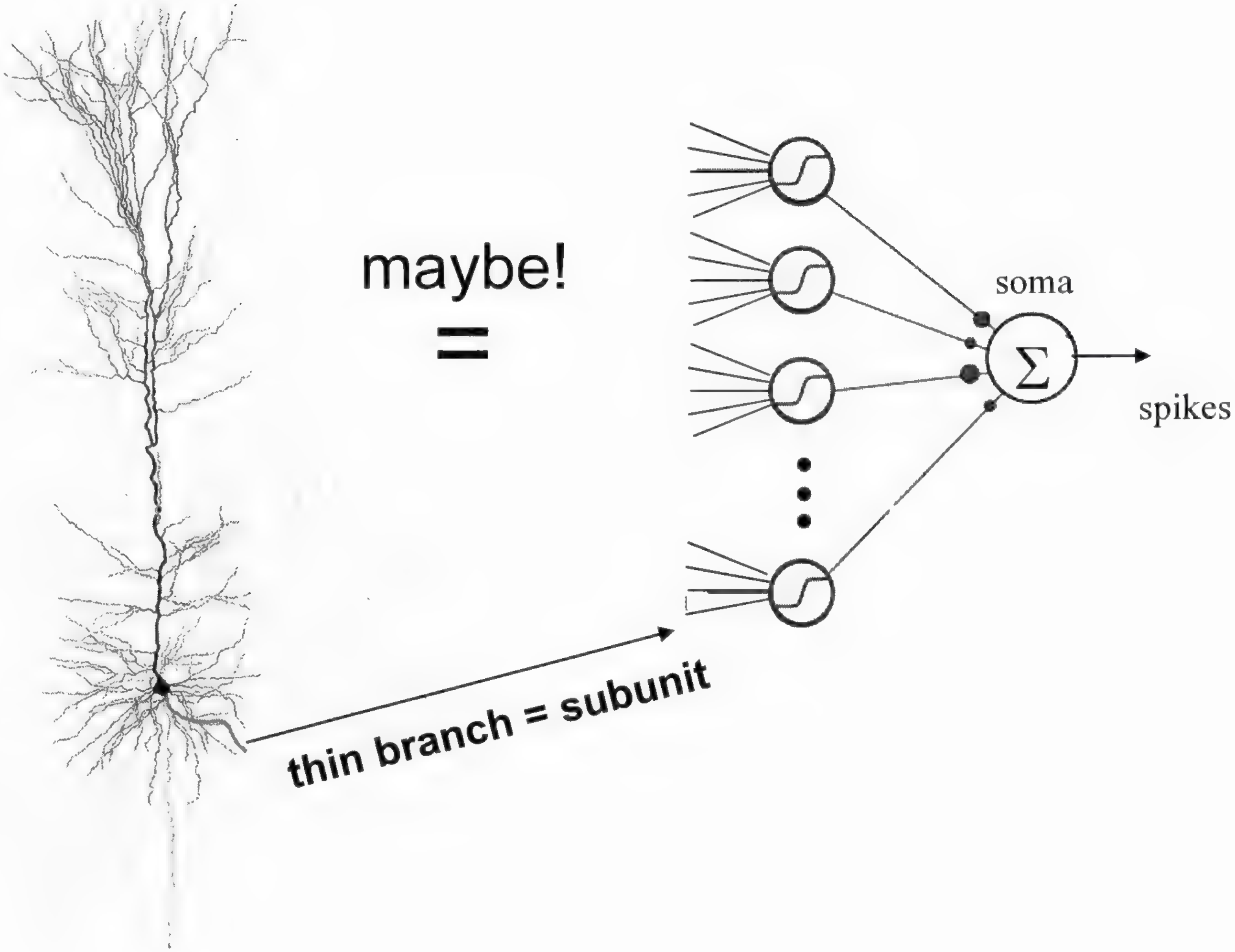

Fully Dispersed
Heterogeneous
Fully Clustered
A
420 µm
B
211 µm
C
soma
2-layer network
w_i
Σ
soma
axon
Point neuron
w_i
Σ
soma
axon
C

Generate 1,000 spatially complex stimulus patterns

The Winner

...and measure firing rate

The 2-layer sum-of-sigmoids model


soma
Σ
spikes

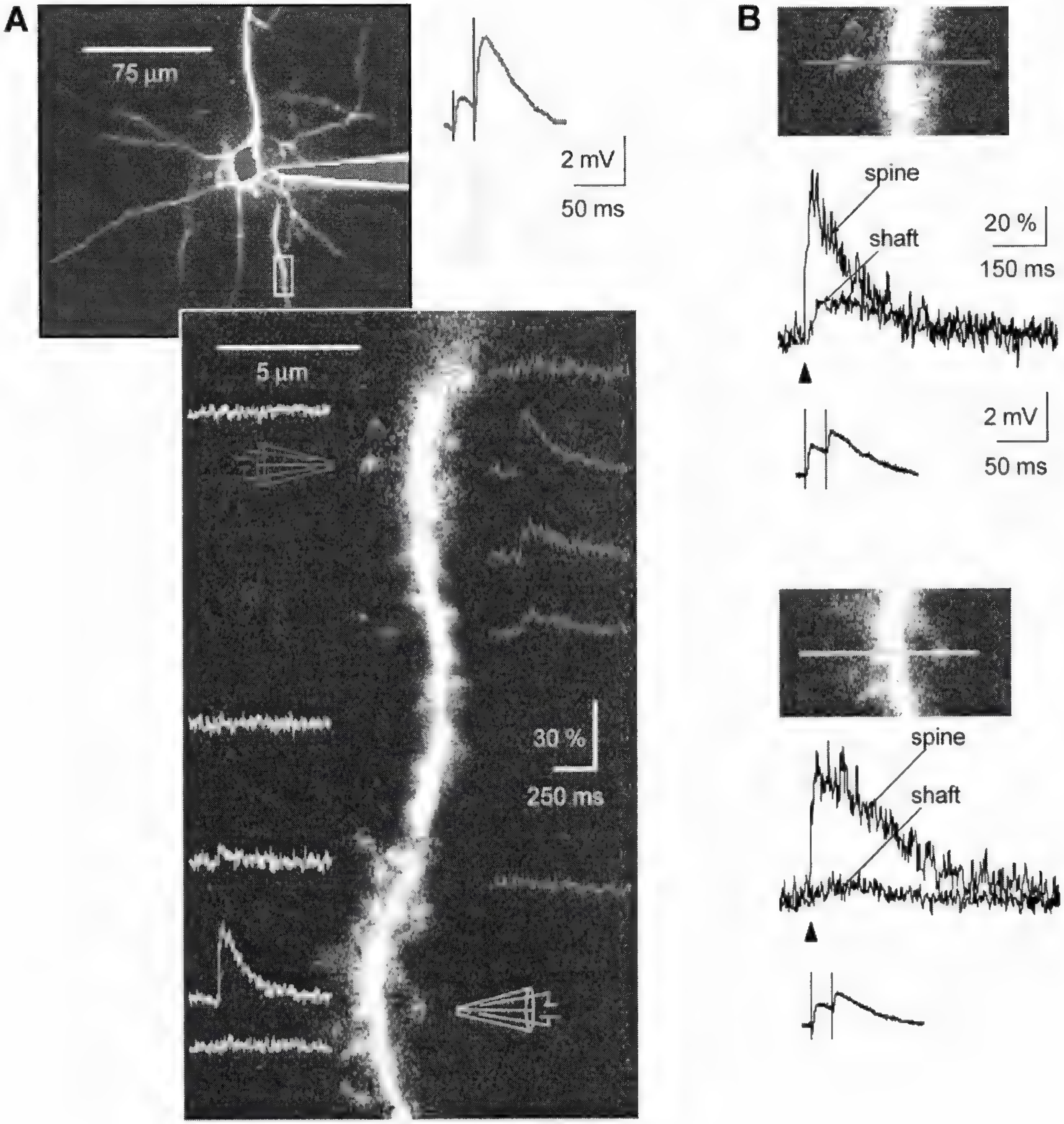
maybe!

=

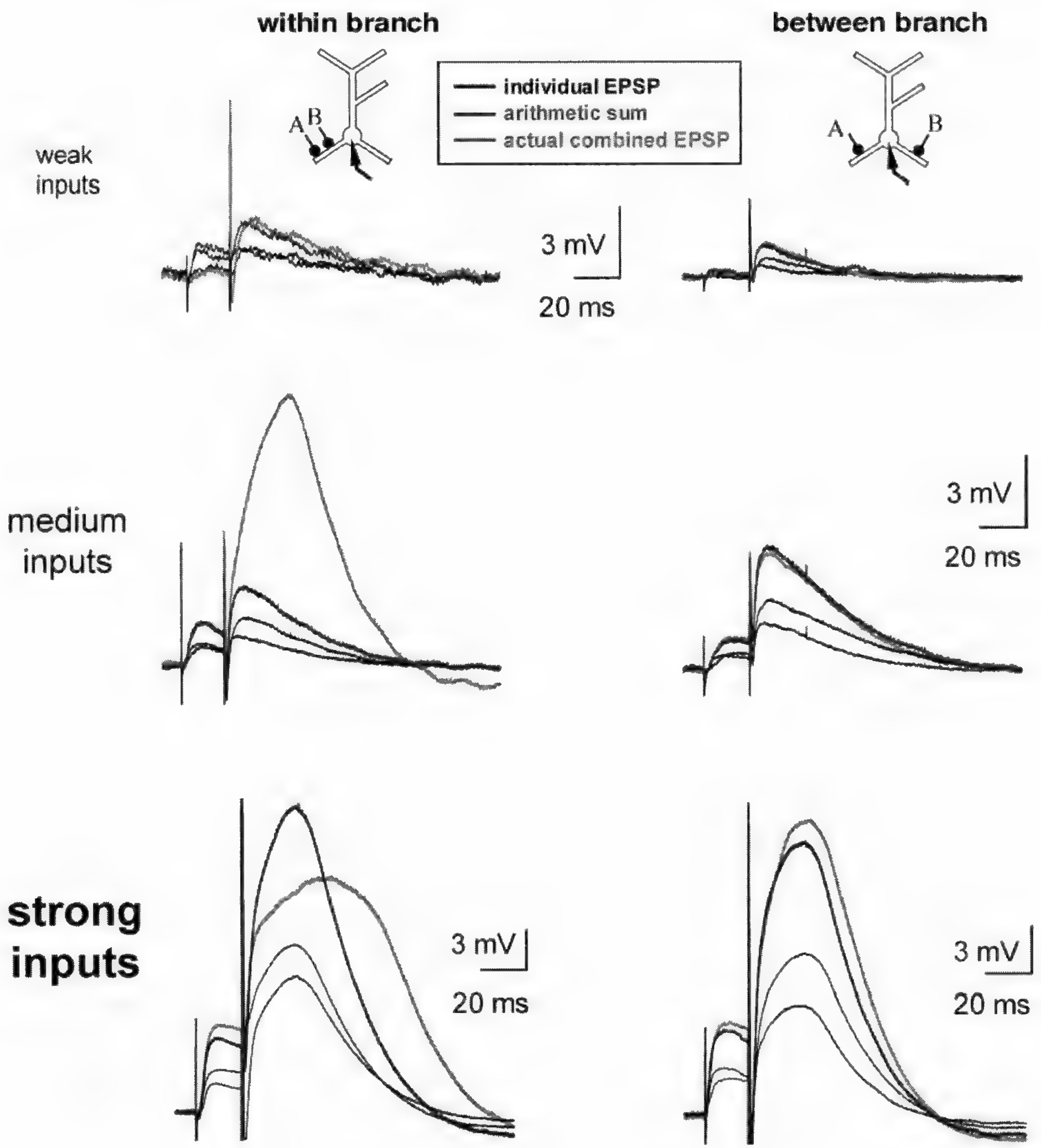
thin branch = subunit

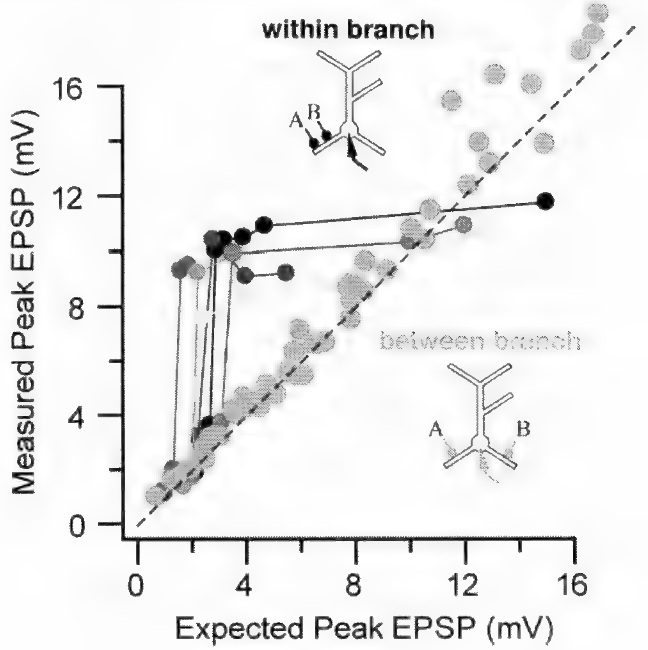
Experimental test in layer 5 pyramidal cells

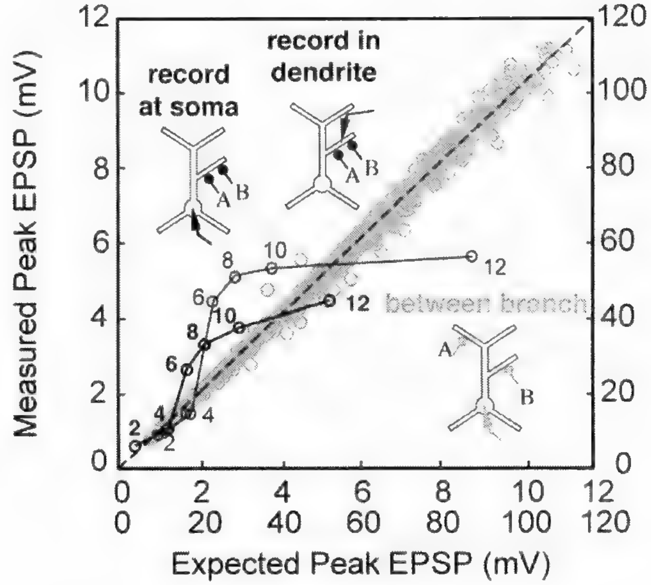
Polsky, Mel & Schiller (2004)


A
75 µm
2 mV
50 ms
5 µm
30 %
250 ms
B
spine
shaft
20 %
150 ms
2 mV
50 ms
spine
shaft

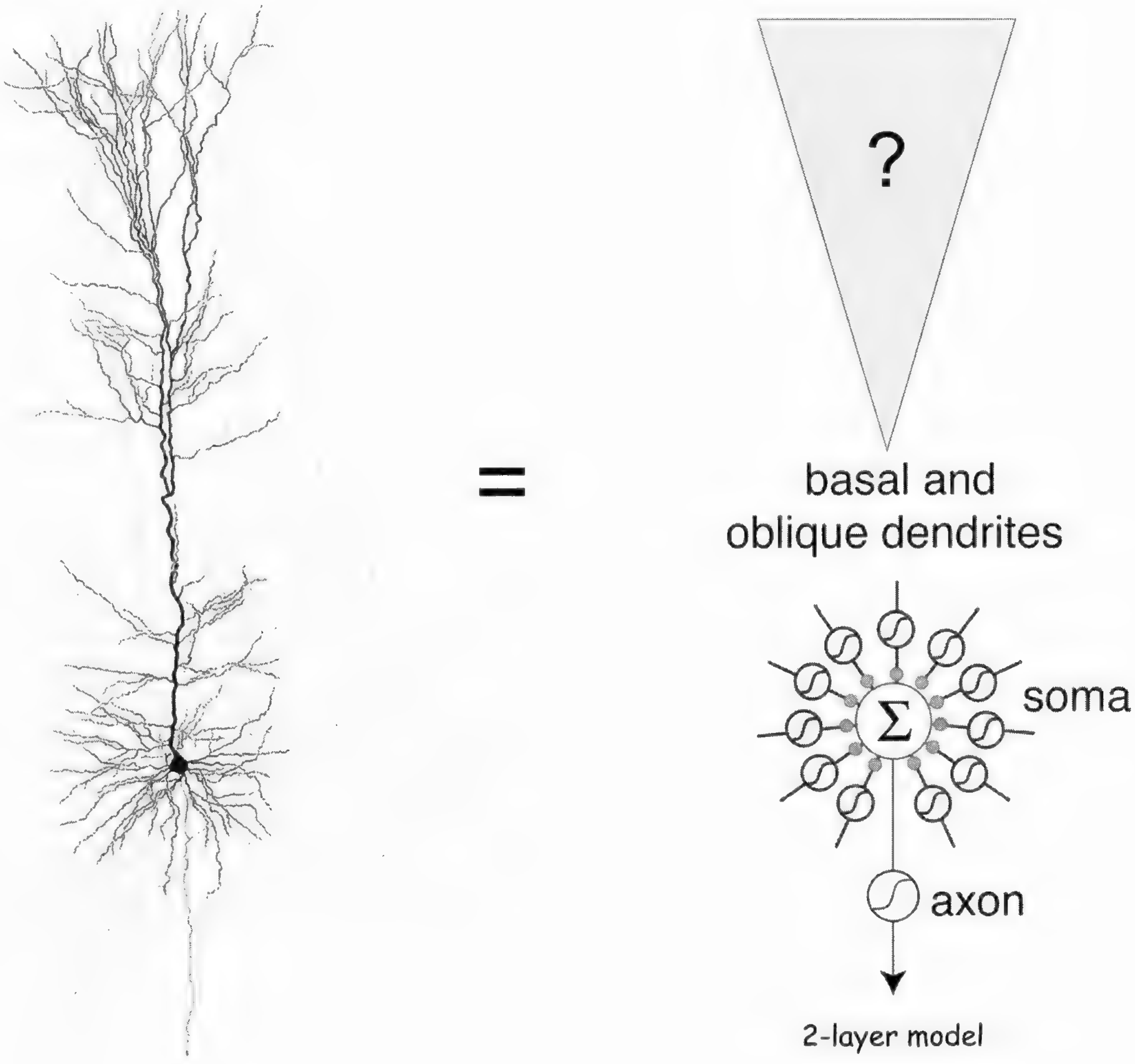
Comparing within- vs. between-branch summation







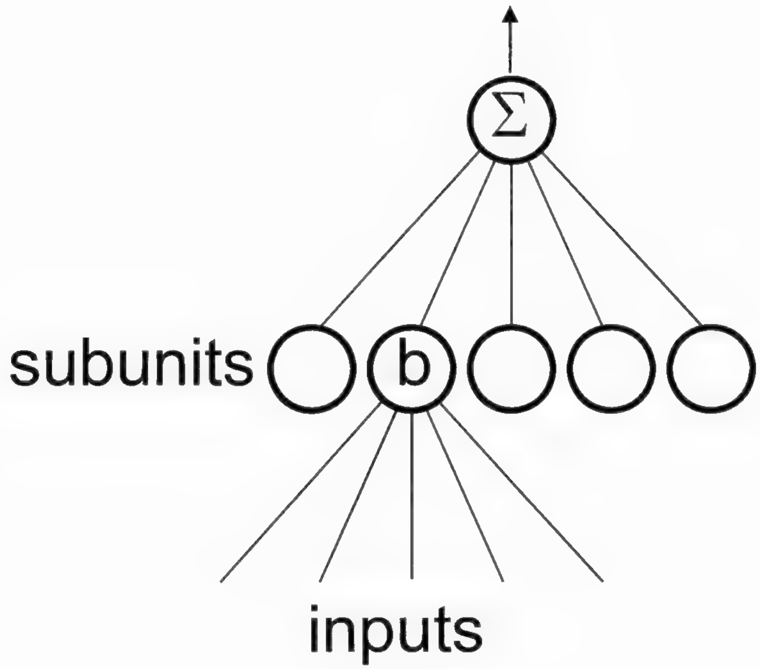
What does a pyramidal neuron do?


?
=
basal and
oblique dendrites
Σ
soma
axon
2-layer model

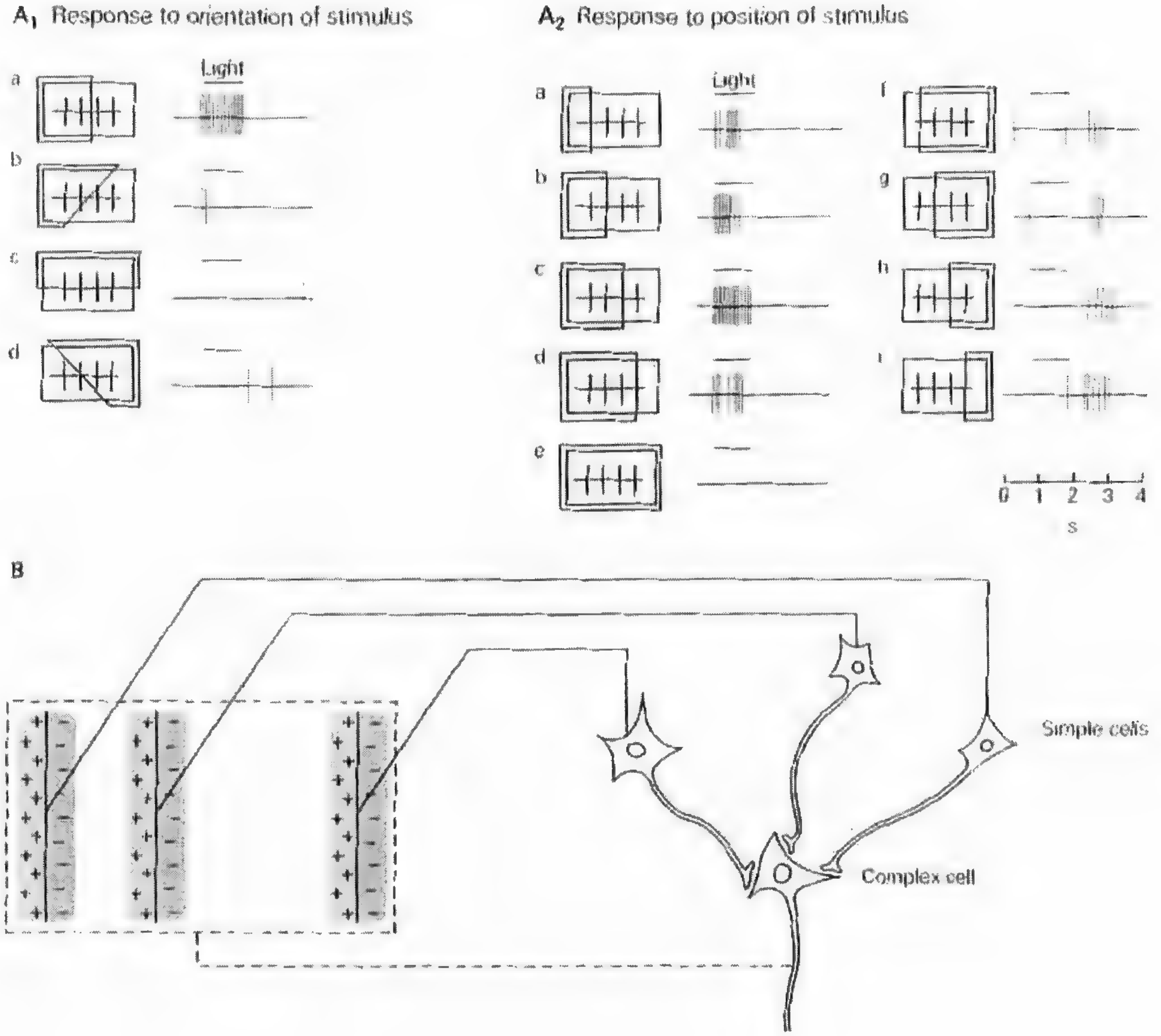
Some "applications" of the 2-layer model

Observation

Many cells in visual cortex have “receptive fields” which can be described as a sum of nonlinear subunits.

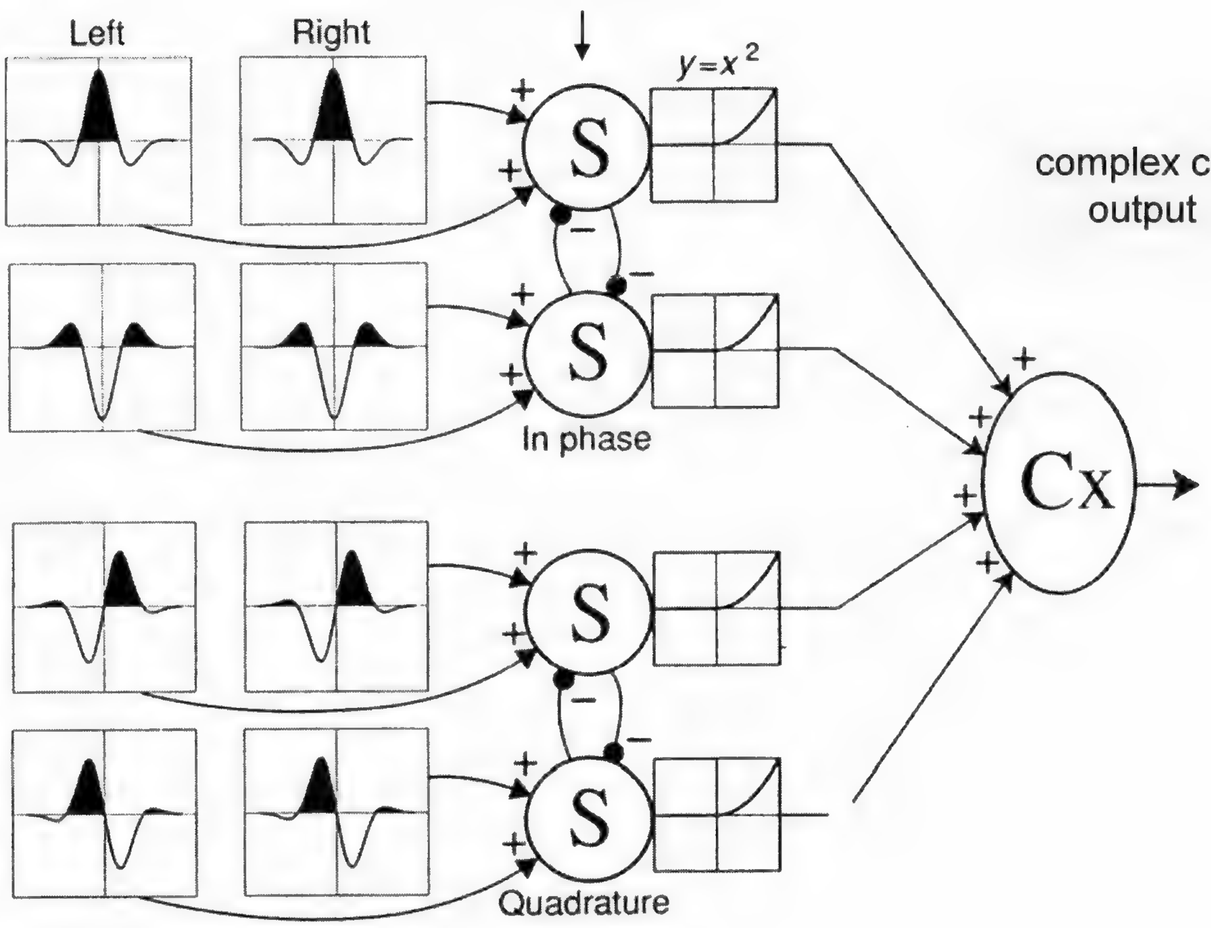

Σ
subunits
b
inputs

Hubel & Wiesel's complex cells


A₁ Response to orientation of stimulus
a
b
c
d
Light
A₂ Response to position of stimulus
a
b
c
d
e
f
g
h
i
Light
0
1
2
3
4
s
B
Simple cells
Complex cell

Binocular Disparity Tuning: The “Energy” Model

Responses of
neurons in
primary visual
cortex to
binocular
stimuli appear
to obey this
mathematical
model...

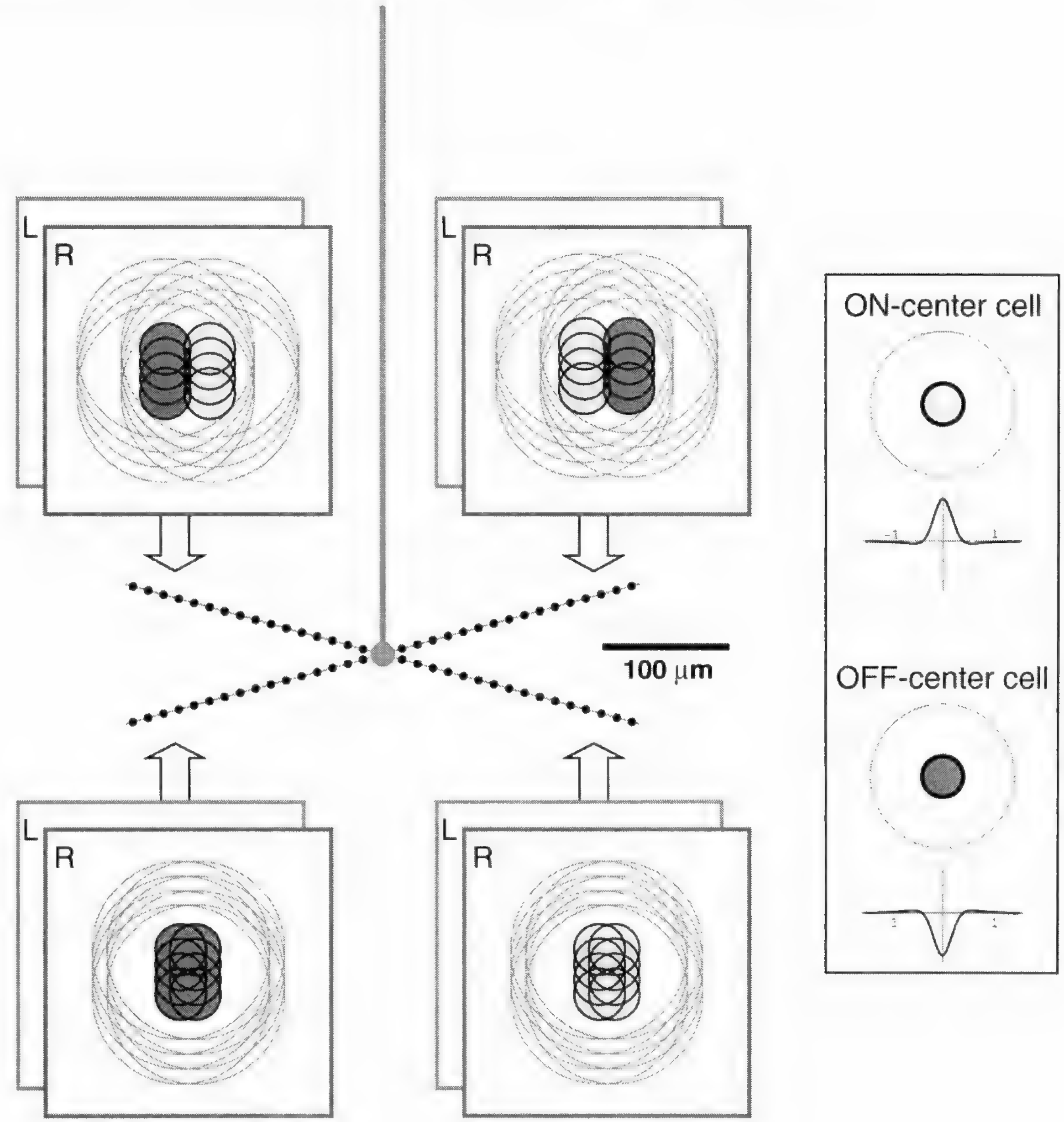

Left
Right
$y=x^2$
S
S
In phase
S
S
Quadrature
+
−
Cx

4 simple cell
modules

complex cell
output

Ohzawa et al. 1997

A dendrite-based binocular complex cell


L
R
L
R
100 μm
L
R
L
R
ON-center cell
-1
1
OFF-center cell
-1
1

4 basal dendrites receive direct LGN input and act like binocular simple cell surrogates...

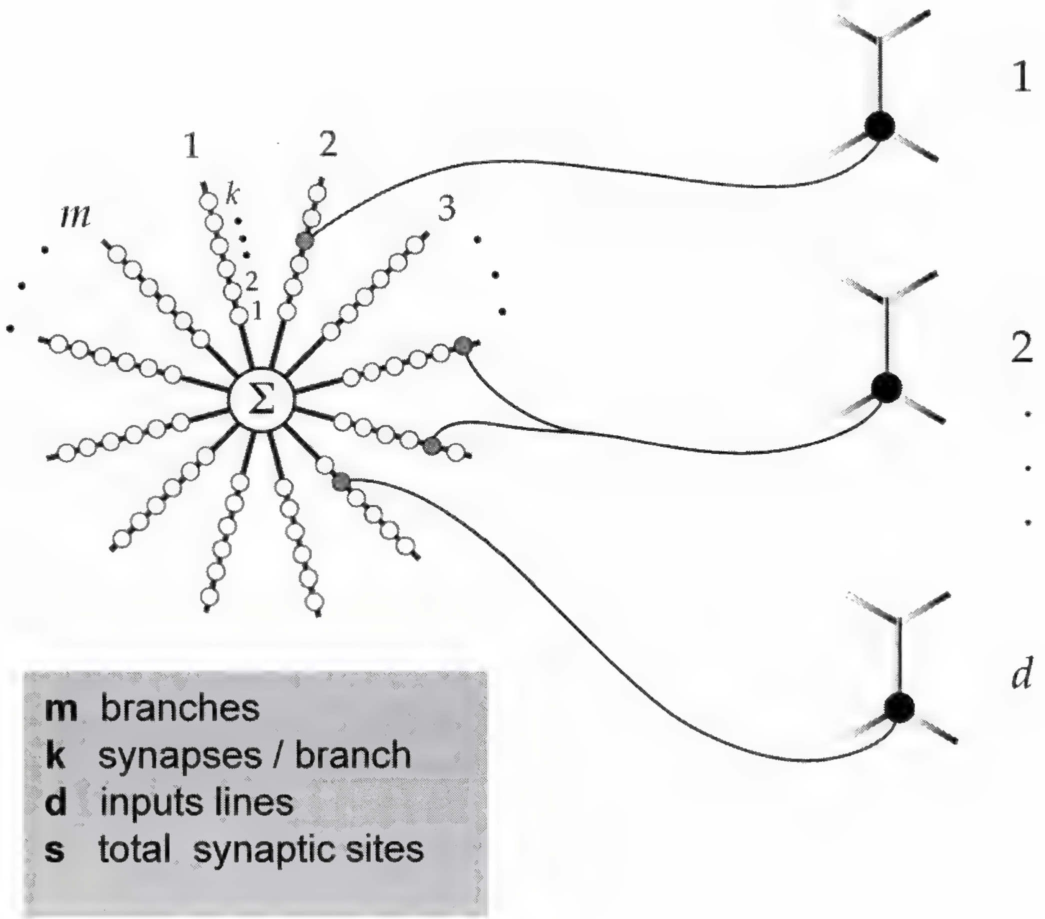
Archie & Mel, 2000

They thought of it first…

One has to consider the possibility that in the monkey the simple-cell step may be skipped, perhaps by summing the inputs from cells in layer 4 on dendrites of complex cells. In such a scheme each main dendritic branch of a complex cell would perform the function of a simple cell.

—David Hubel, *Nature* 299, 515-524, 1982.

Dendritic subunits increases memory capacity



m branches
k synapses / branch
d inputs lines
s total synaptic sites

Linear Model (bits)

$$B_L = 2\log_2 \binom{s+d-1}{s}$$

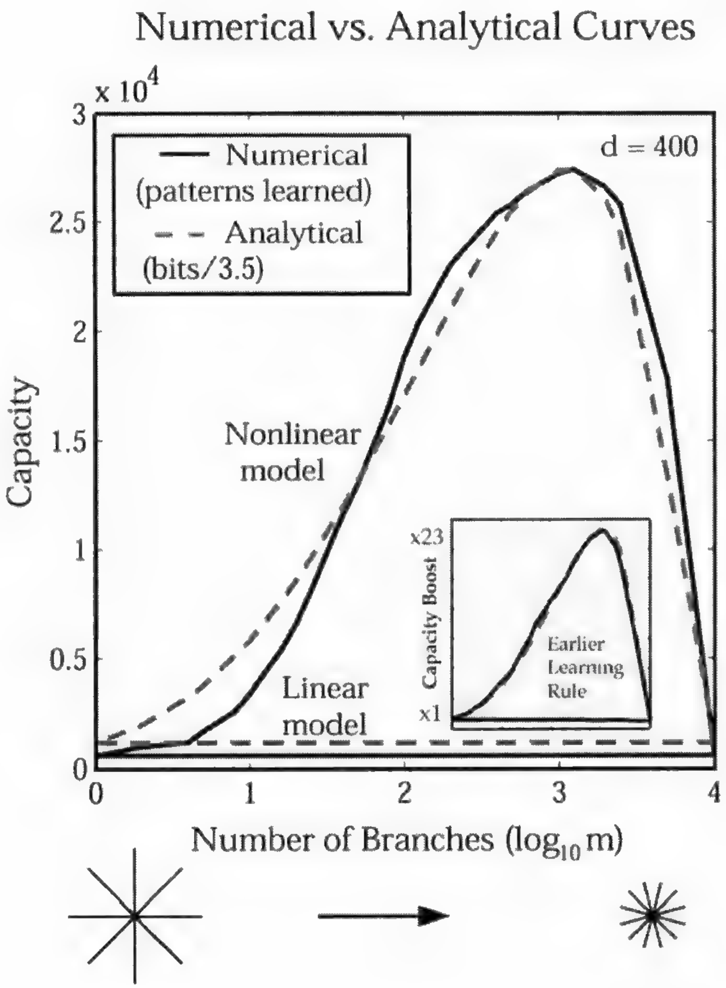
Nonlinear Model (bits)

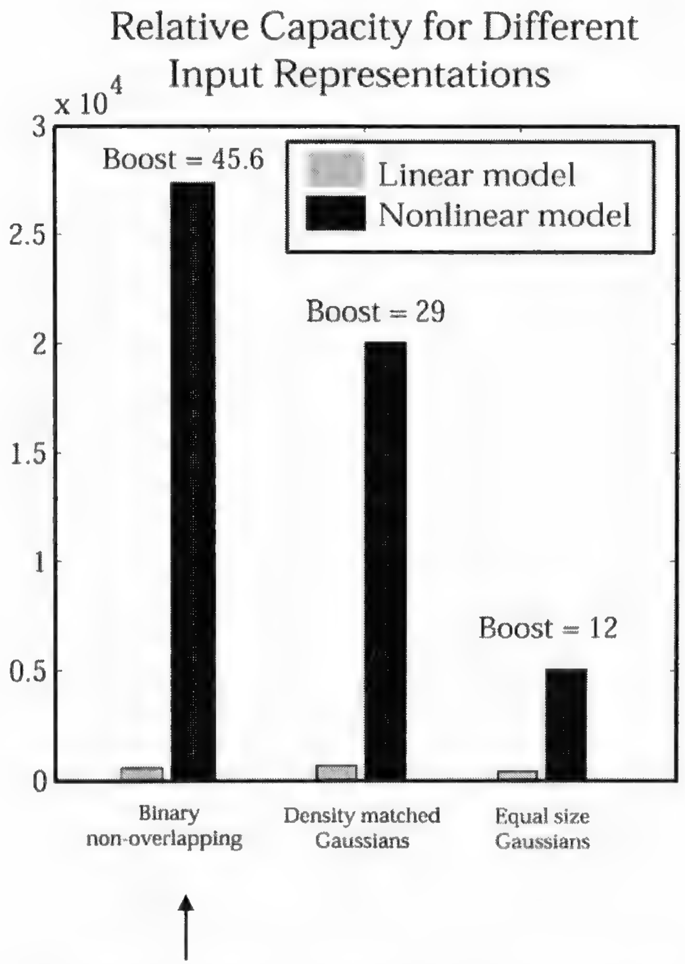
$$B_N = 2\log_2 \binom{\binom{k+d-1}{k}+m-1}{m}$$

Poirazi & Mel (2001)

Analytical vs. empirical capacity measures

In actual learning trials, empirical capacity for cells of different shapes and sizes is close to analytical prediction.

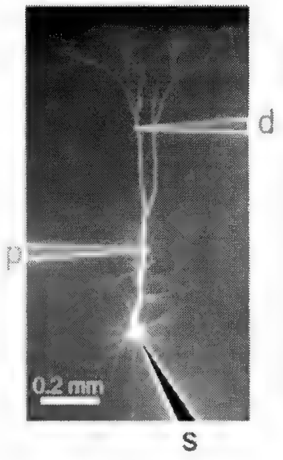


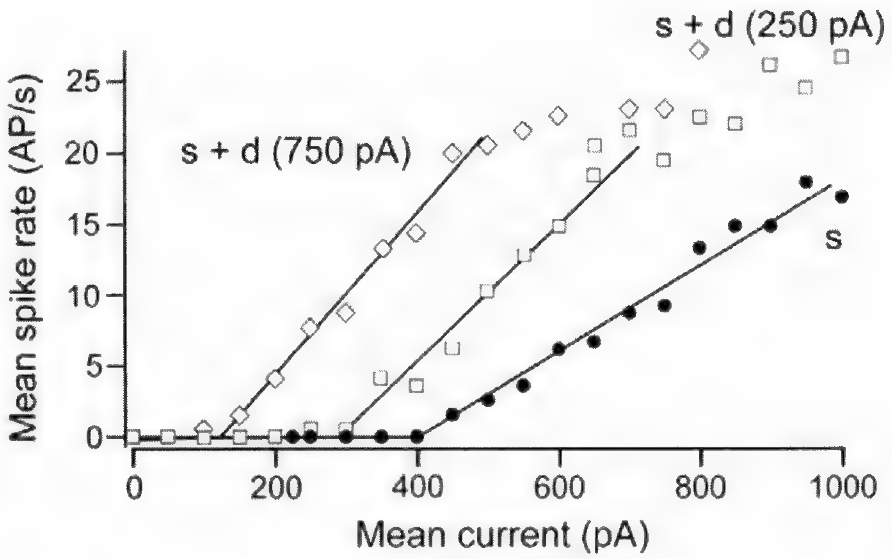


Empirical capacity boost for nonlinear subunit-containing cells is larger than analytical prediction.

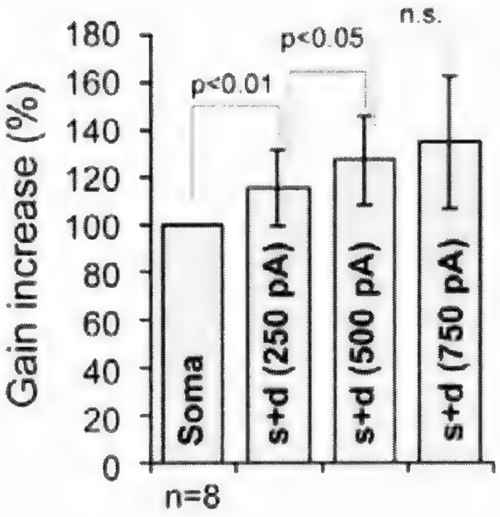
Some refinements…

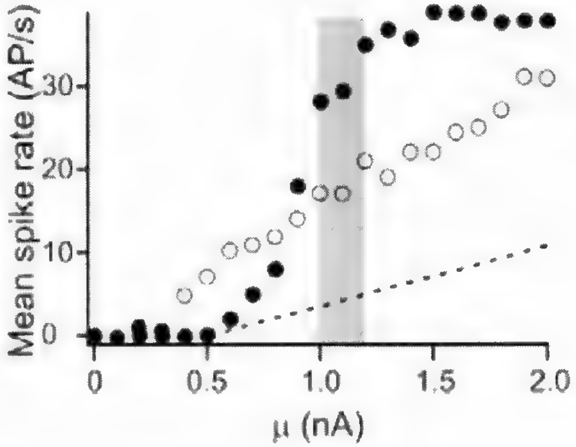
How does the apical tuft fit in?





Apical input increases gain

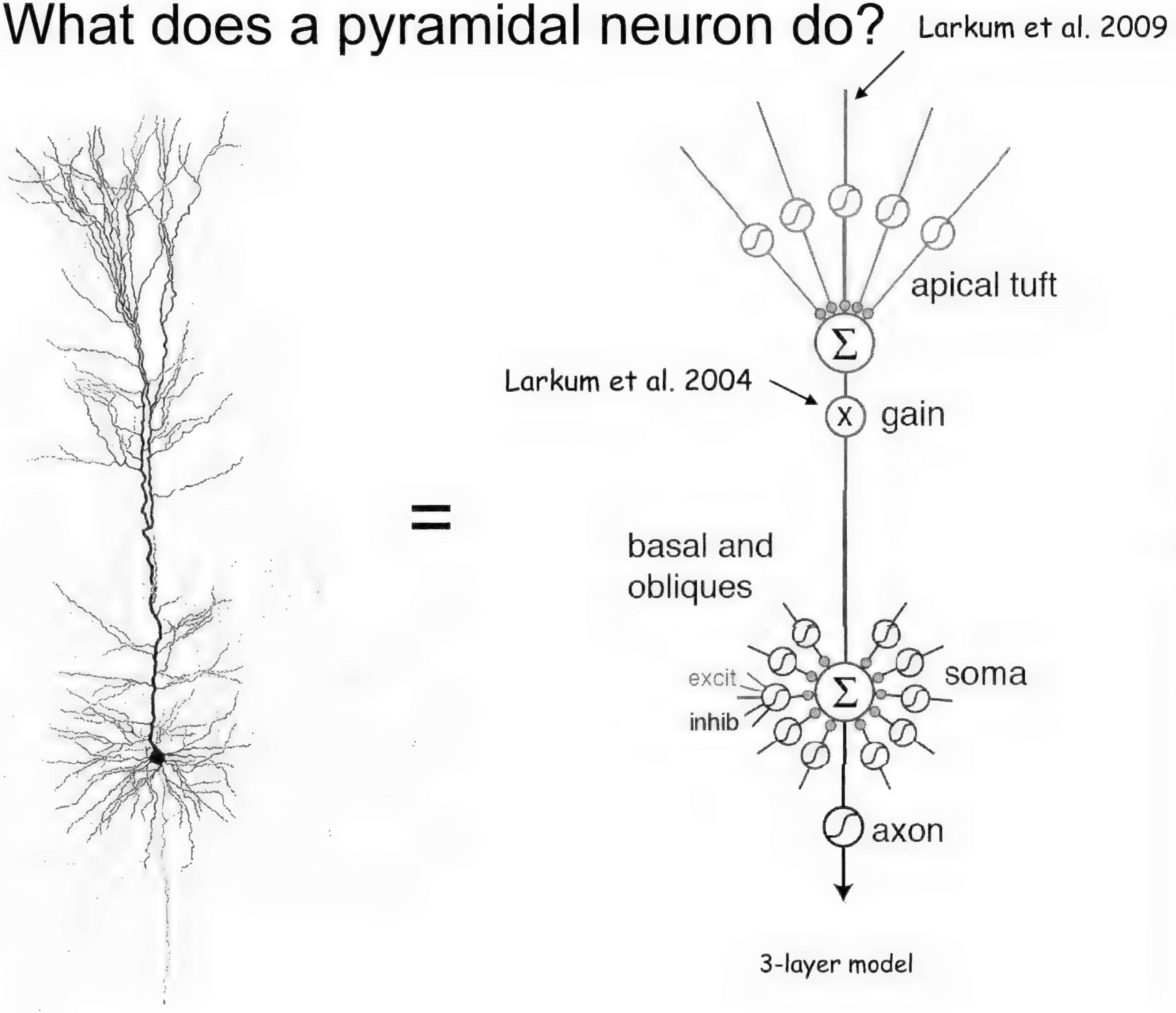




Blocking Ca++ reduces gain

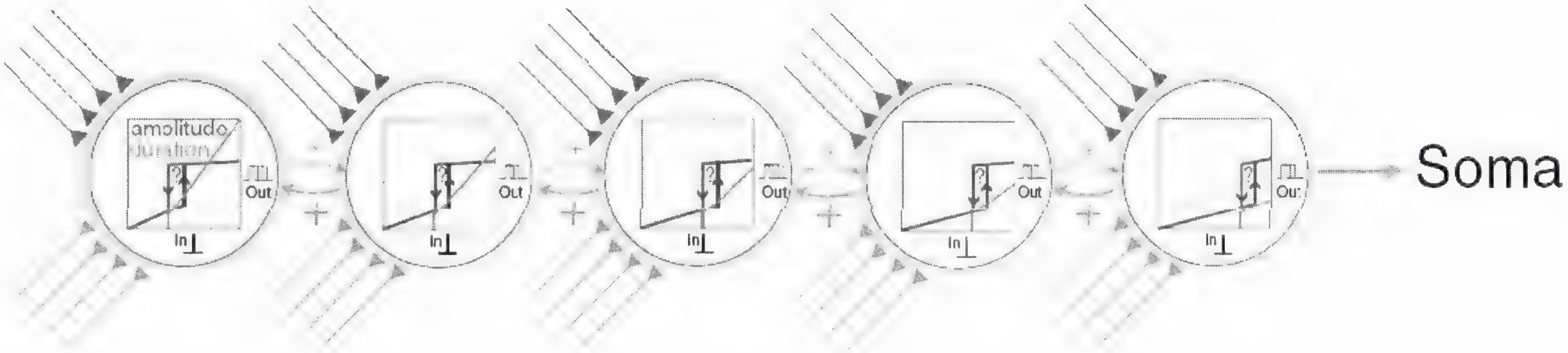
Larkum, Senn & Luscher (2004)

What does a pyramidal neuron do?


Larkum et al. 2009
=
apical tuft
Σ
Larkum et al. 2004
x
gain
basal and obliques
excit
inhib
Σ
soma
axon

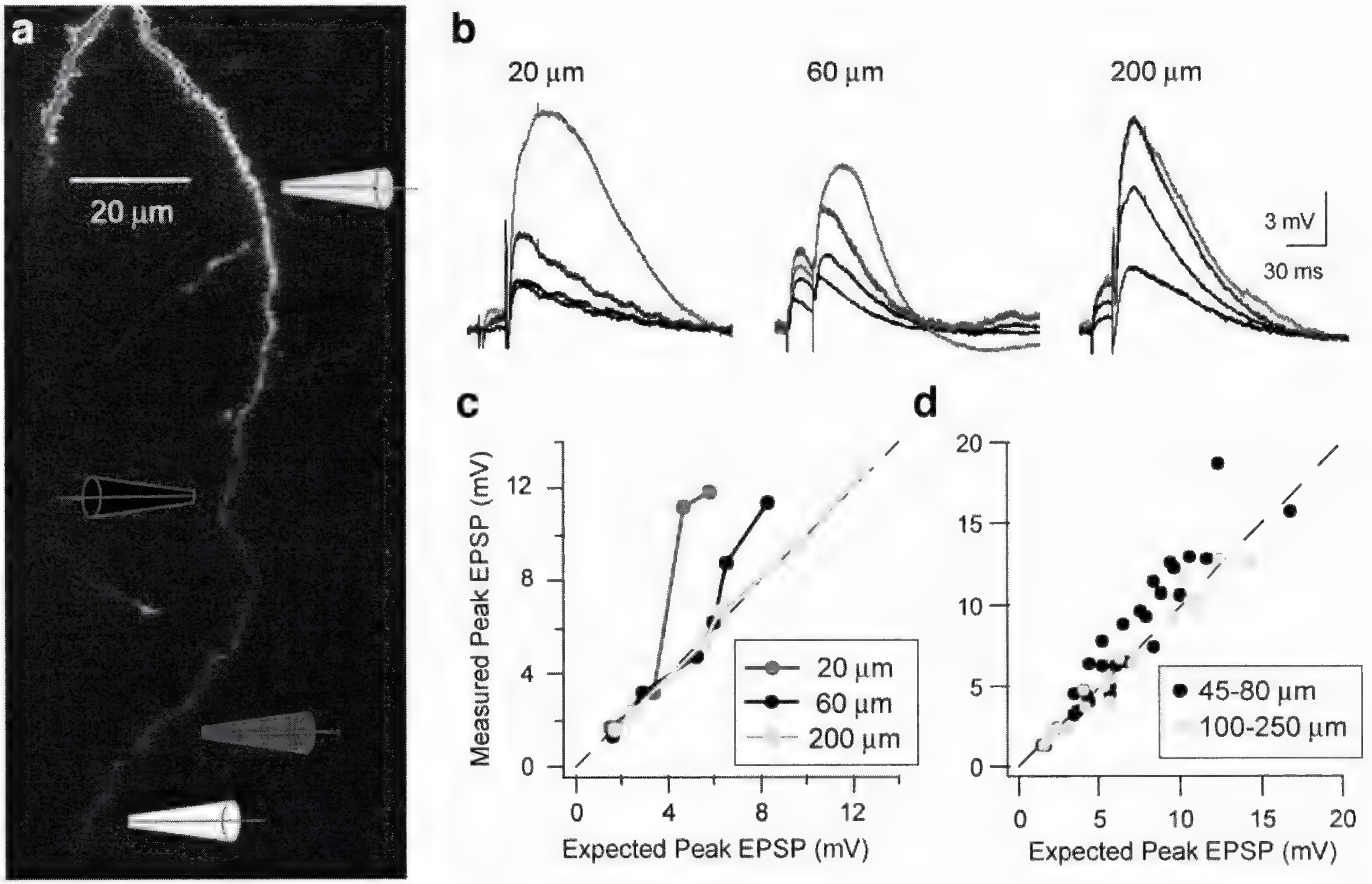
3-layer model

Are thin branches really monolithic subunits?


amplitude
duration
?
Out
In
+
-
Soma

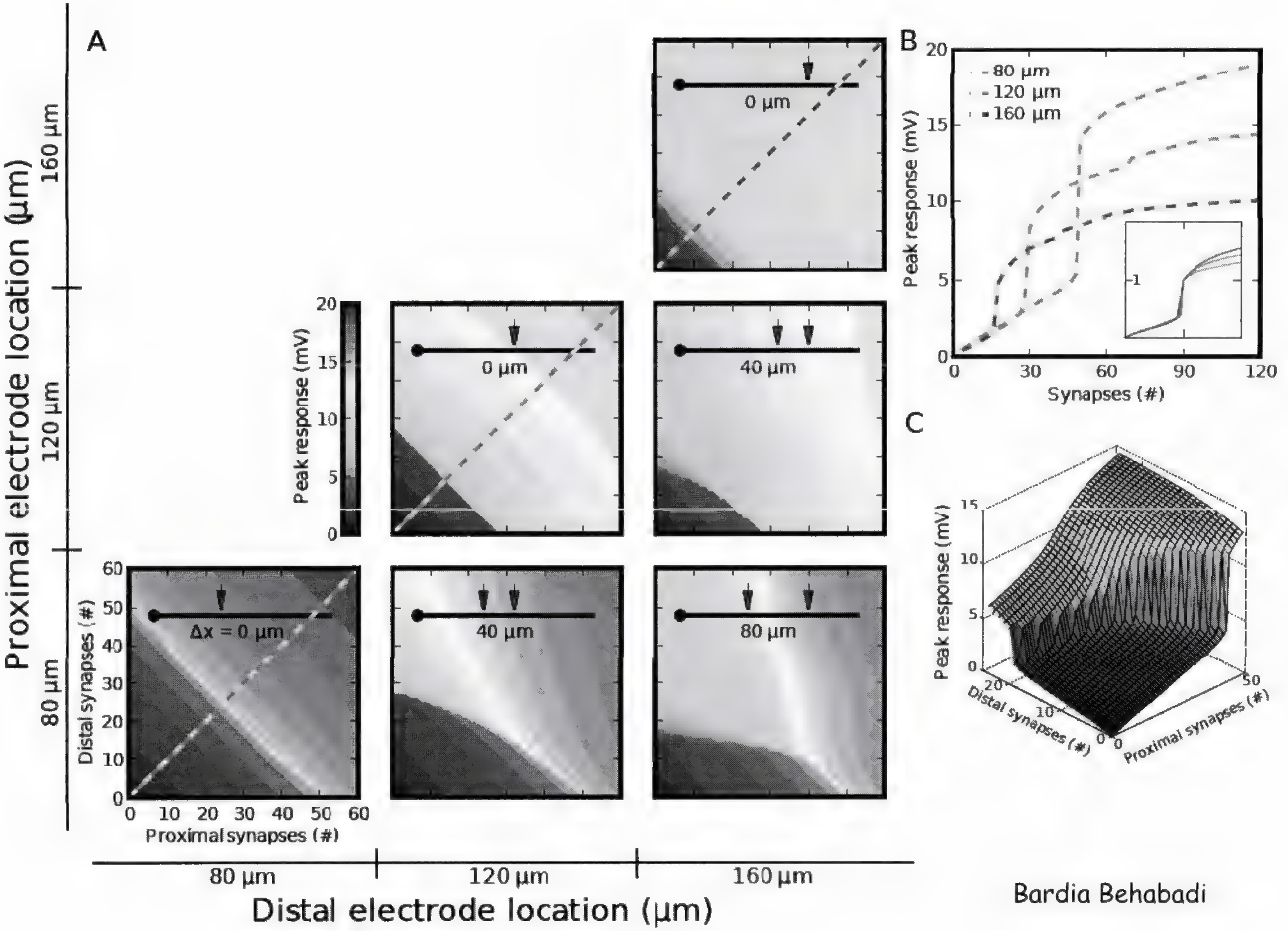
Major et al. (2008)

Are thin branches really monolithic subunits?



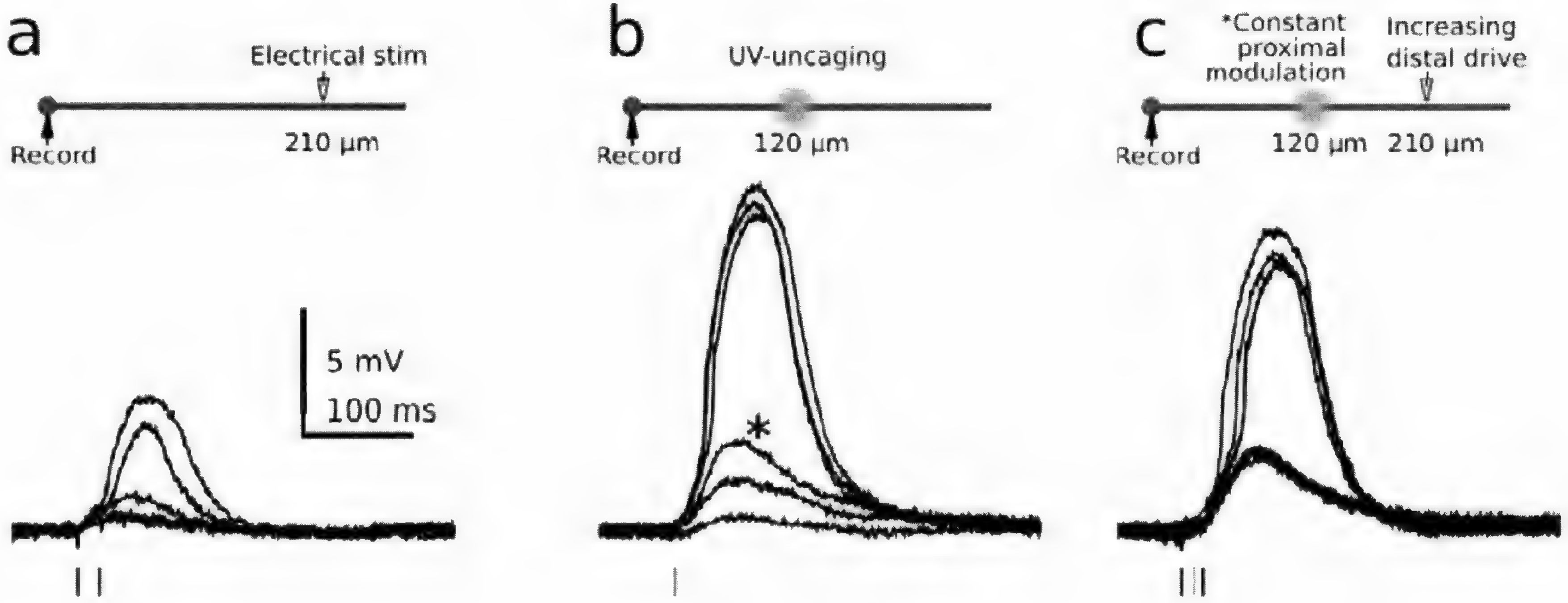
Polsky et al. (2004)

Exploring the 2-input-output behavior of a thin dendrite



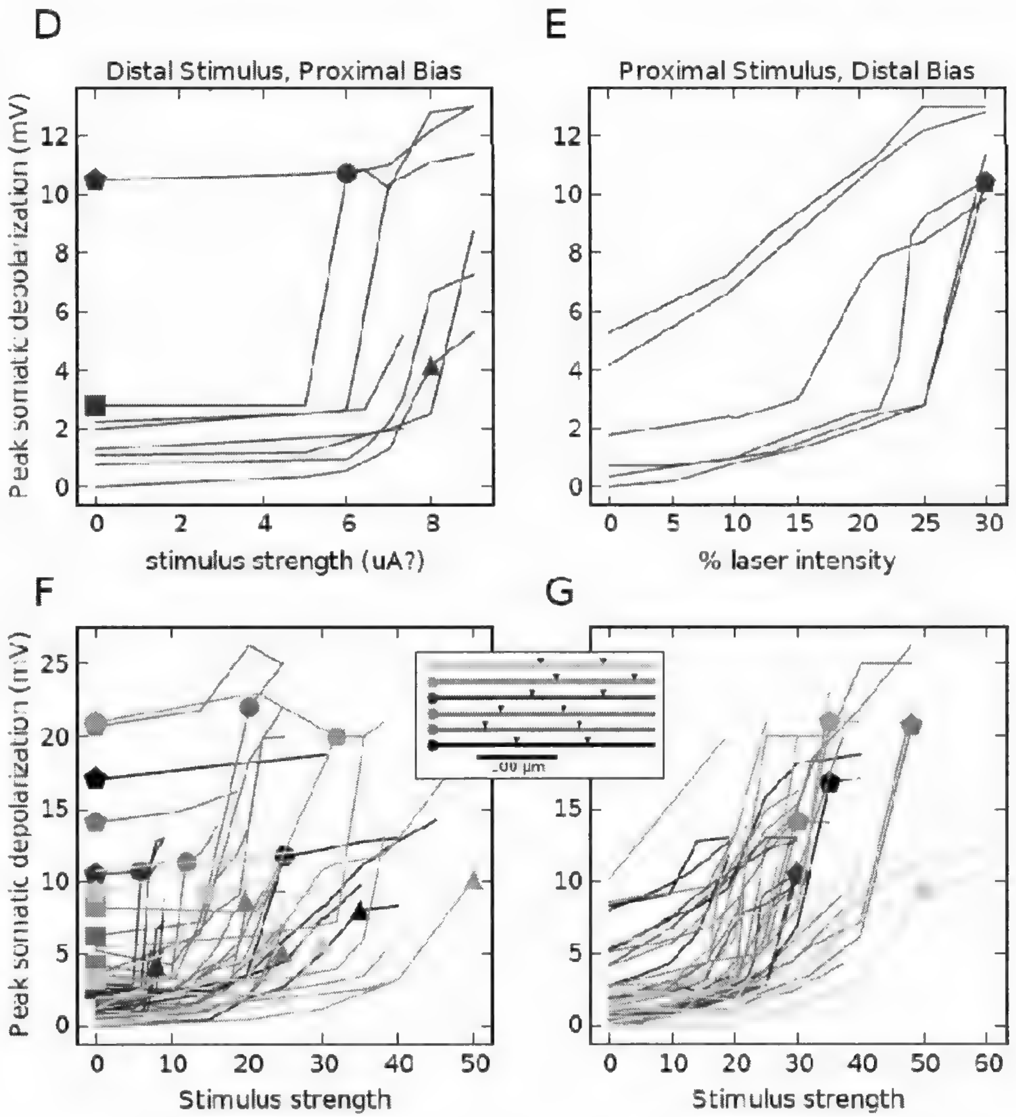
Bardia Behabadi

Experimental test of 2-input summation

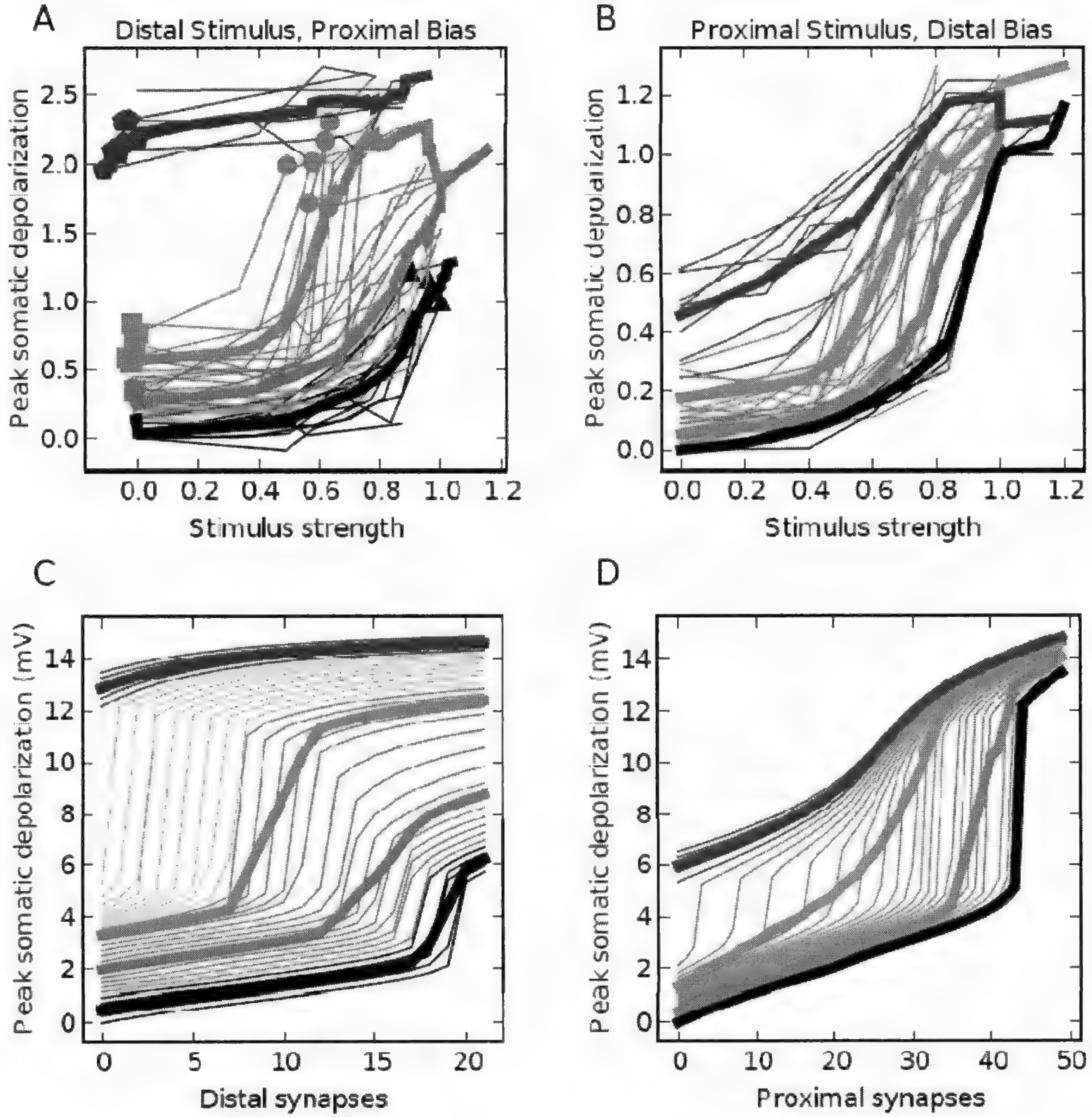


w/ Alon Polsky and Jackie Schiller

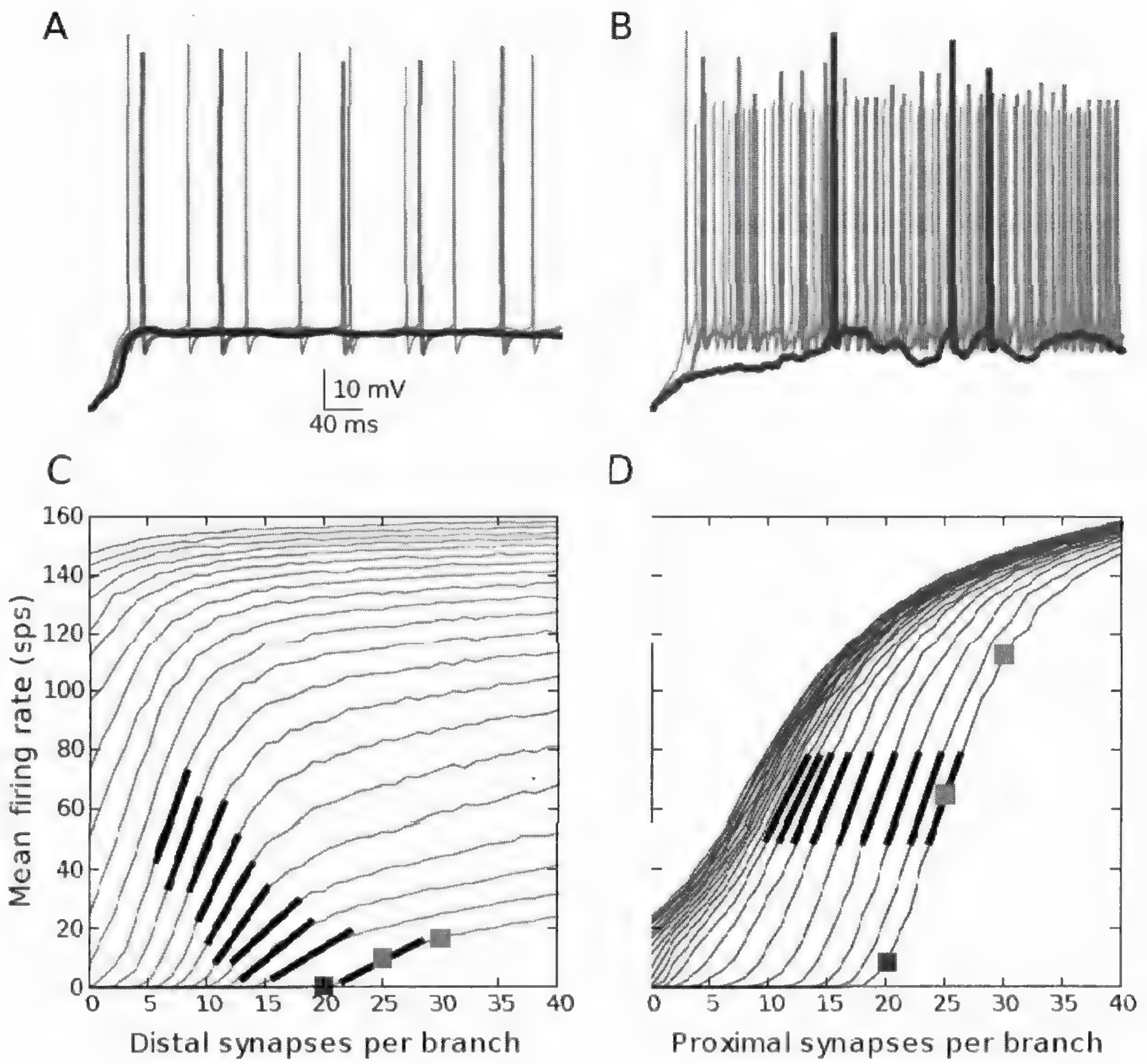
Experimental test of 2-input summation



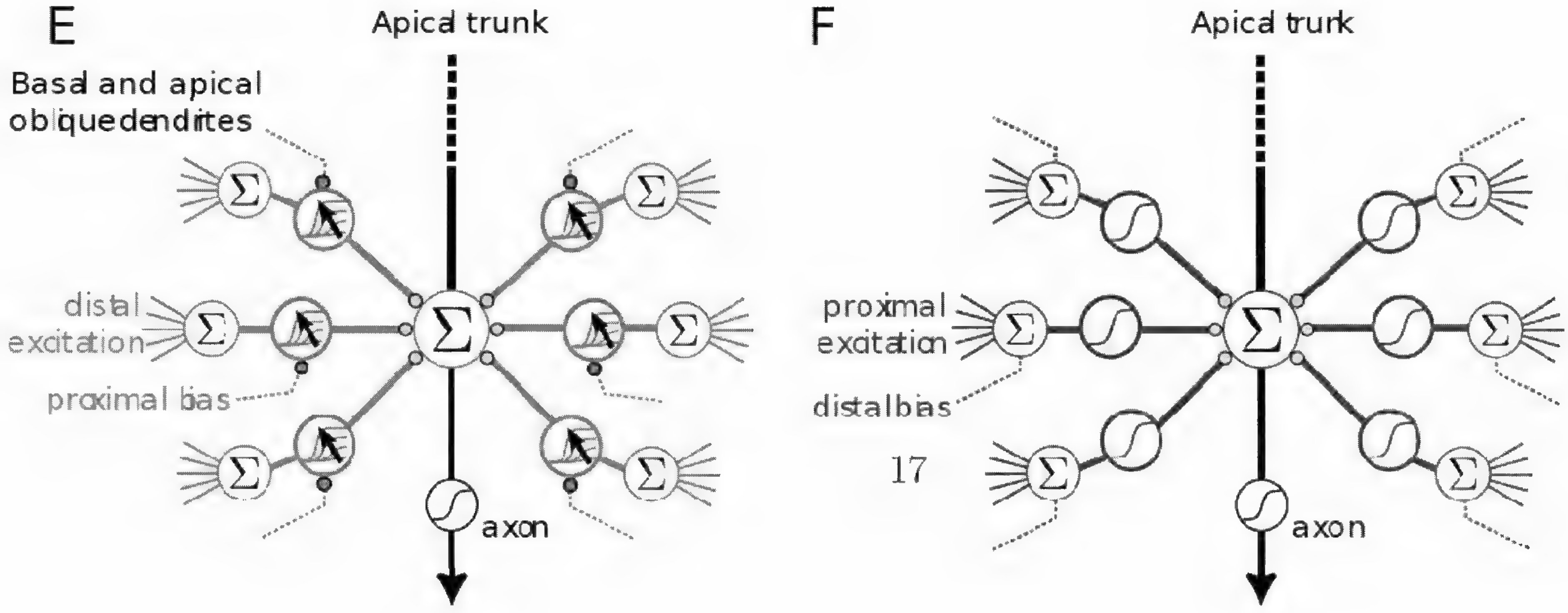
Comparing model and data

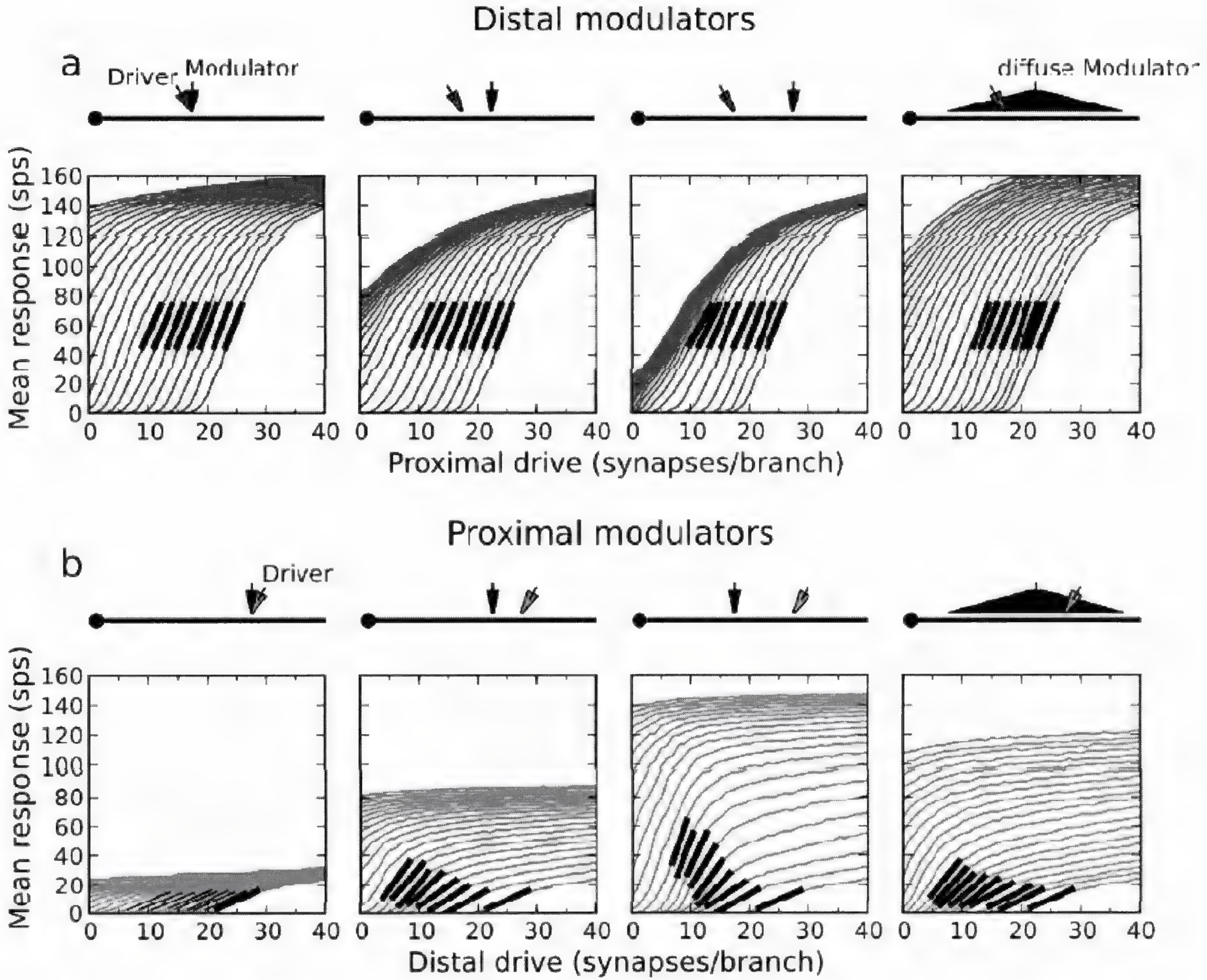


Prediction: similar pattern in spike-rate regime

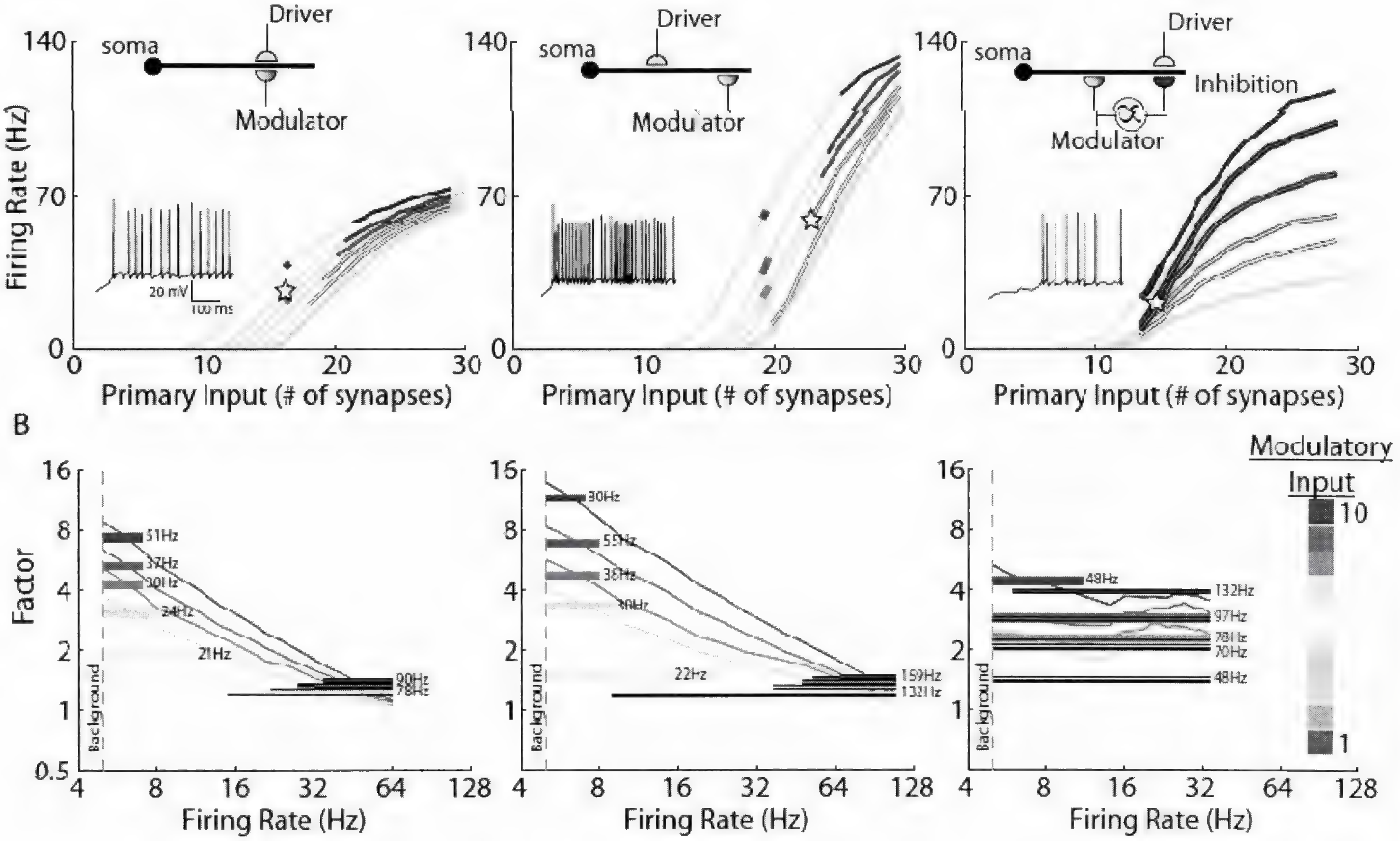


An elaborated model of basal dendrites with contextual modulation

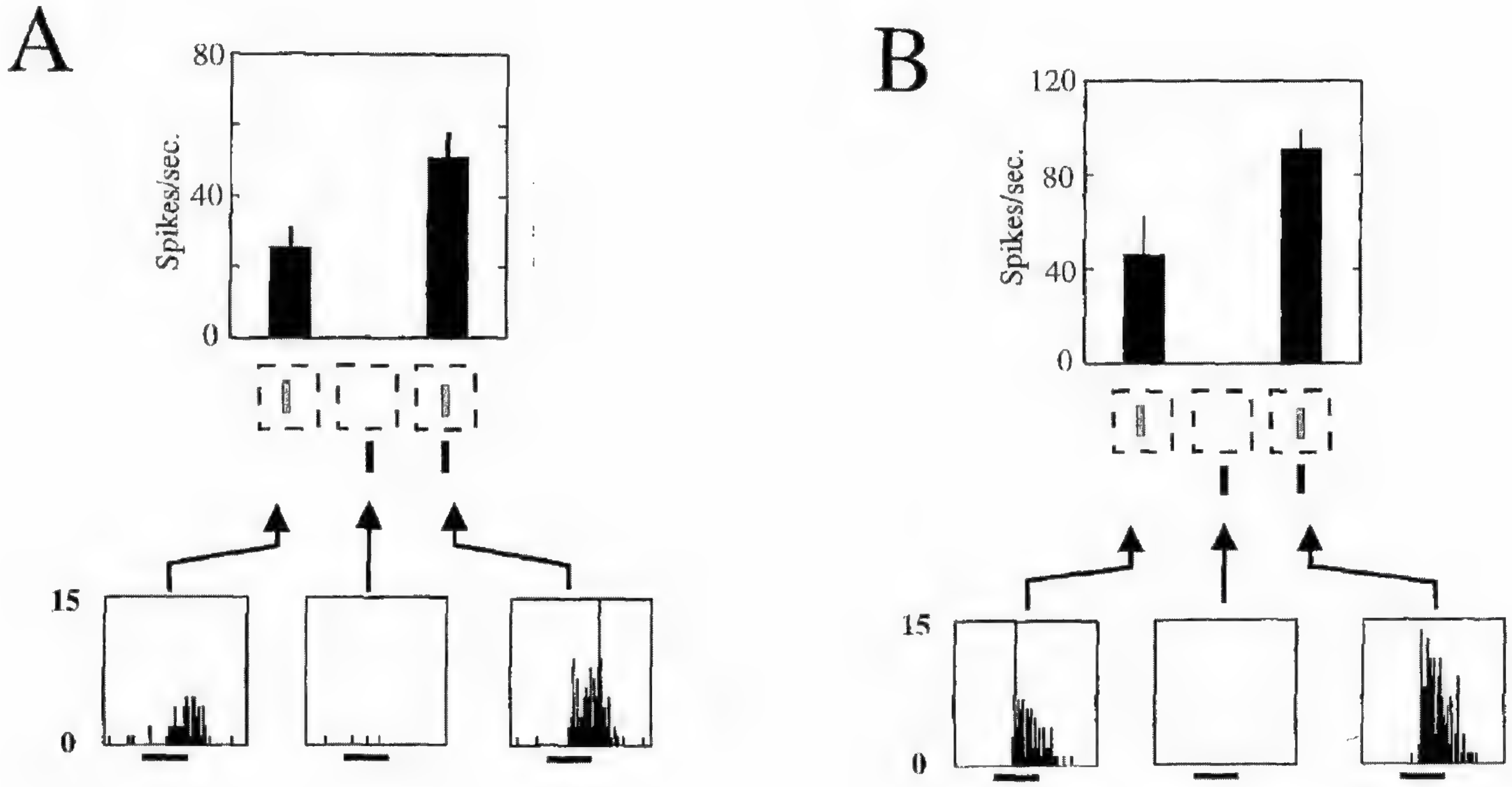

E
Apical trunk
Basal and apical oblique dendrites
distal excitation
proximal bias
axon
F
Apical trunk
proximal excitation
distal bias
axon


Distal modulators
a
Driver
Modulator
diffuse Modulato
Mean response (sps)
160
140
120
100
80
60
40
20
0
0
10
20
30
40
Proximal drive (synapses/branch)
Proximal modulators
b
Driver
Mean response (sps)
Distal drive (synapses/branch)

"Scaling competence"

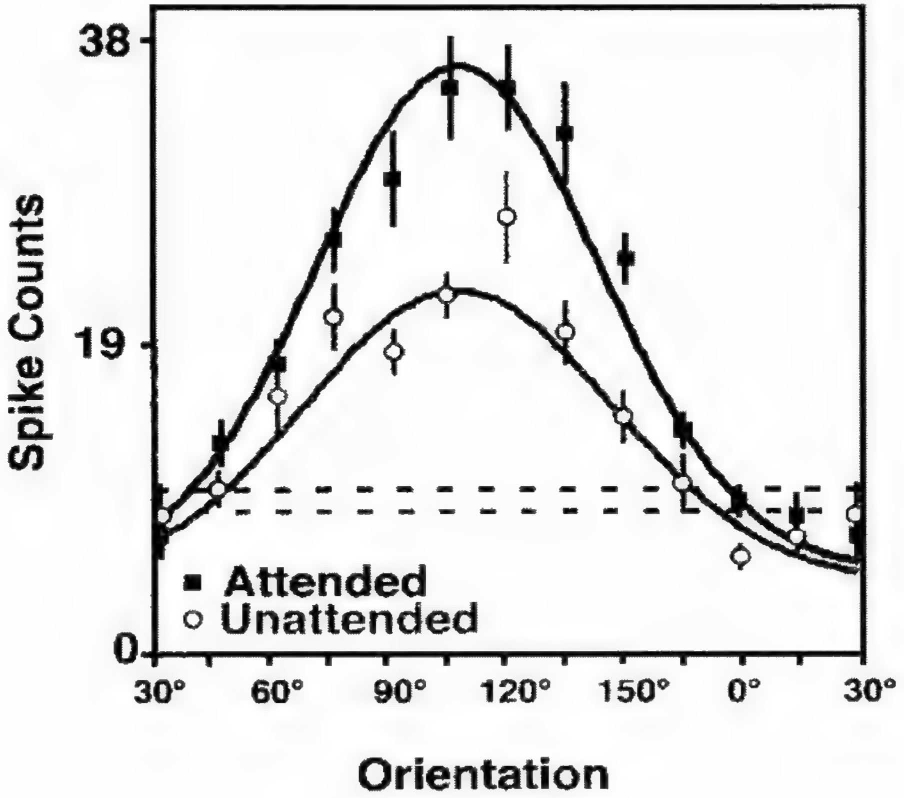


Contextual boosting in monkey V1



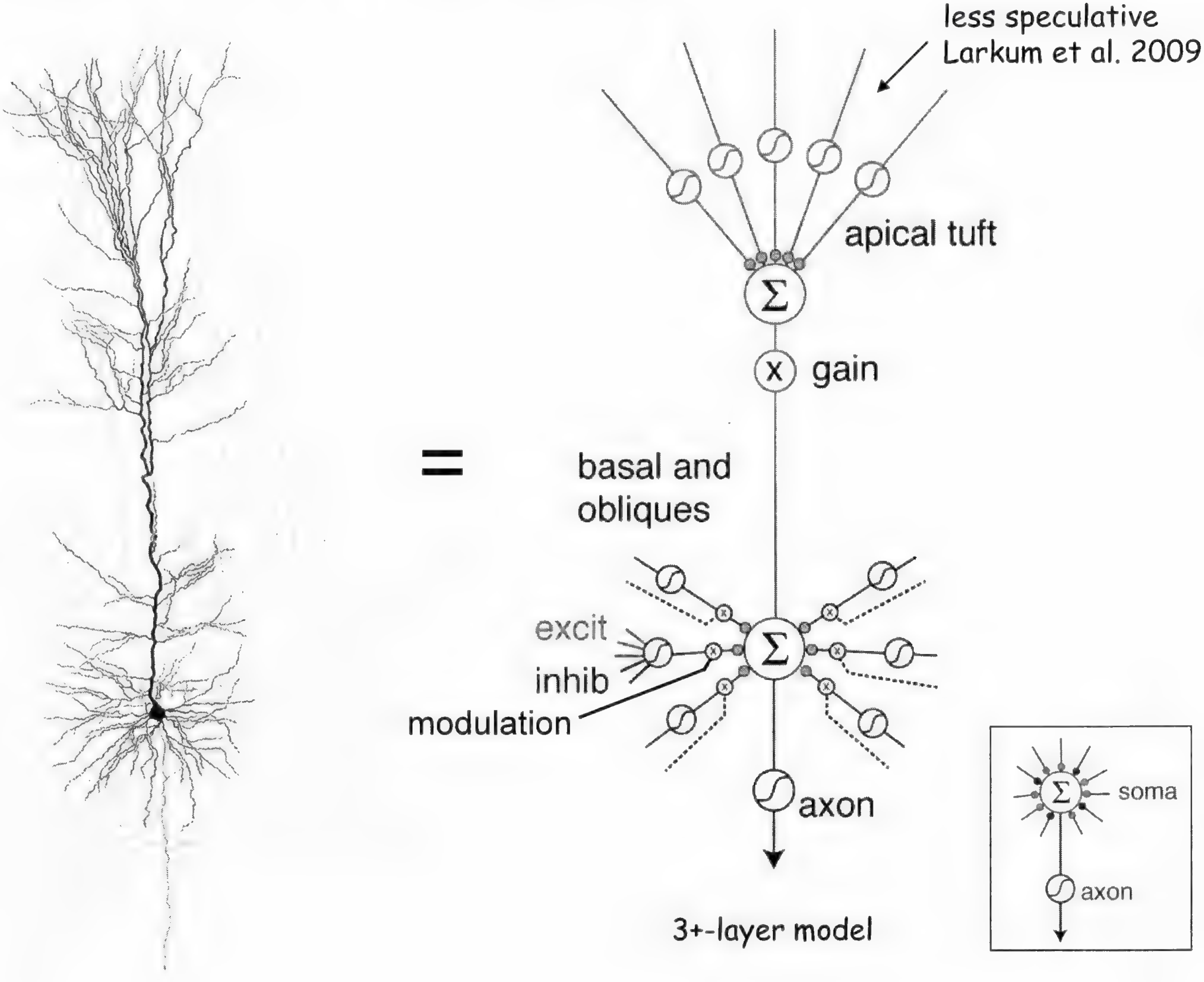
Kapadia et al. (1995)

Attention also multiplies



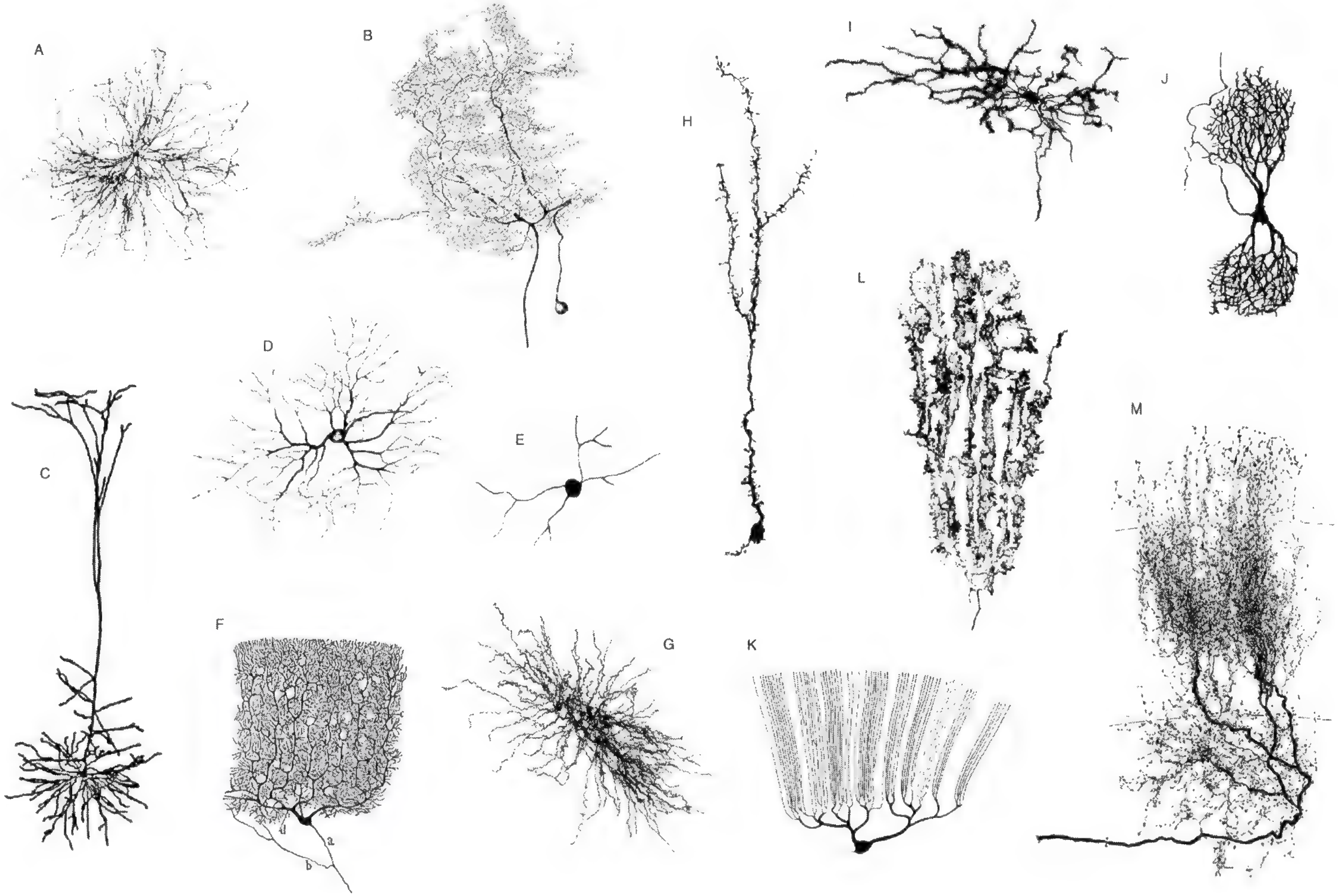
McAdams & Maunsell (1999)

What does a pyramidal neuron do?

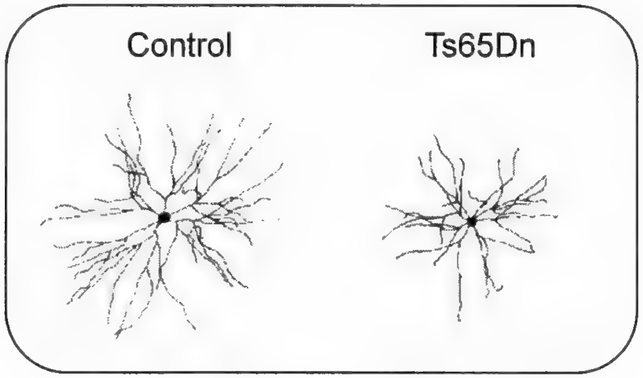

less speculative
Larkum et al. 2009
apical tuft
Σ
x
gain
=
basal and
obliques
excit
inhib
modulation
Σ
axon
3+-layer model
Σ
soma
axon

Open questions

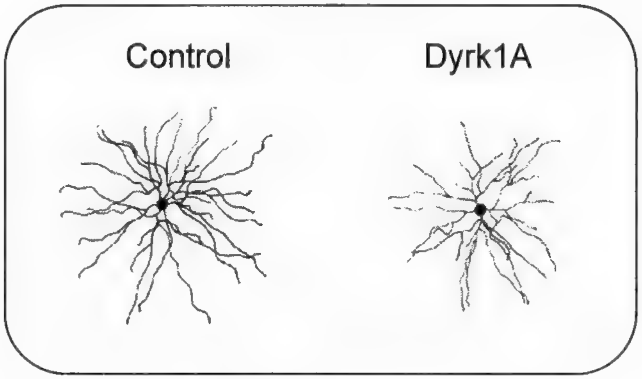
How many different models do we need?


A
B
H
I
J
L
D
E
C
M
F
G
K
d
a
b

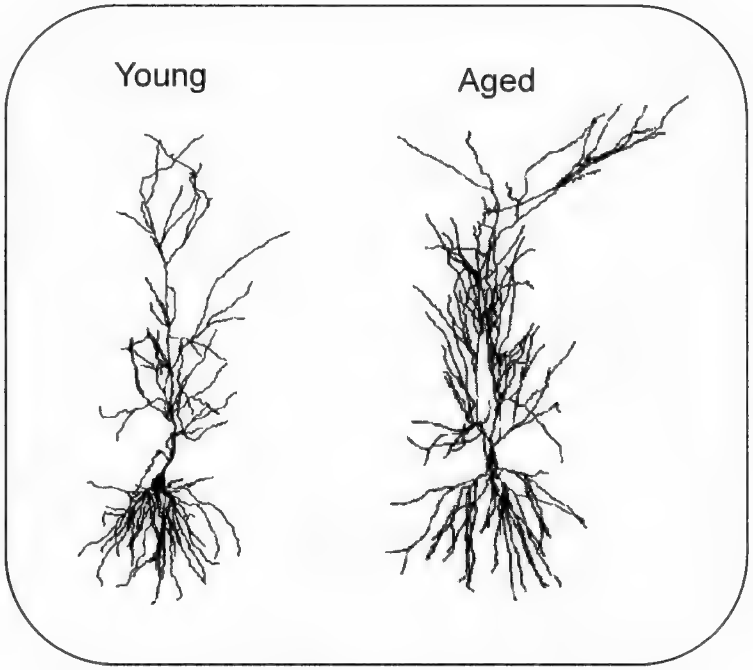
Morphological changes: which ones matter most?


Control
Ts65Dn

Frontal cortex
Dierssen et al. (2003)


Control
Dyrk1A

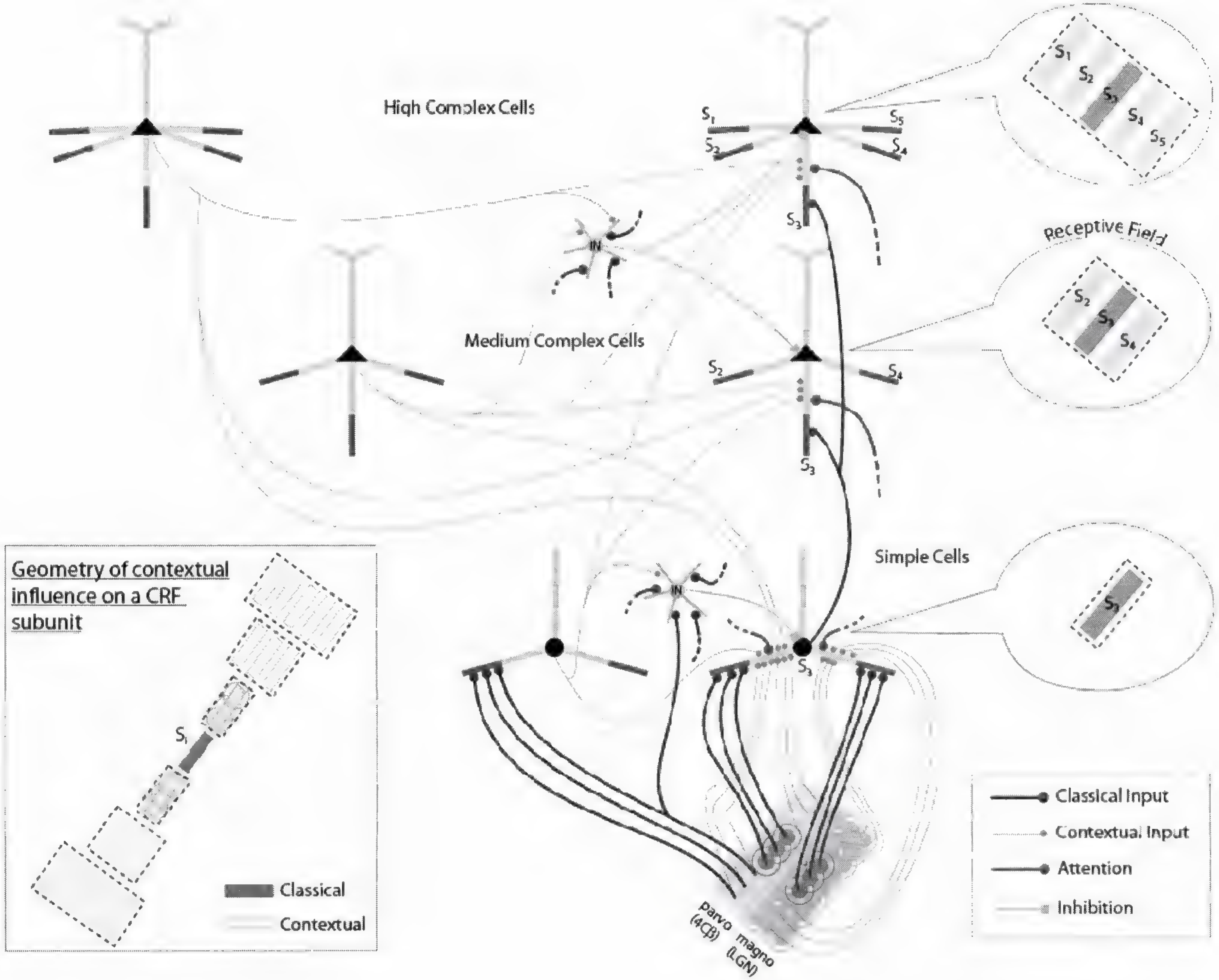
Motor cortex
Benavides-Piccione et al. (2005)



CA1
Pyapali & Turner (1996)

200 μm

How do dendrites contribute to neural circuits?


High Complex Cells
S1
S2
S3
S4
S5
Receptive Field
S2
S3
S4
Medium Complex Cells
IN
S2
S4
S3
Simple Cells
S3
IN
Geometry of contextual influence on a CRF subunit
S1
Classical
Contextual
parvo (4Cβ)
magno (LGN)
Classical Input
Contextual Input
Attention
Inhibition

Dendritic Computation

Current lab members

Bardia Behabadi
Monika Jadi

Lab alumni

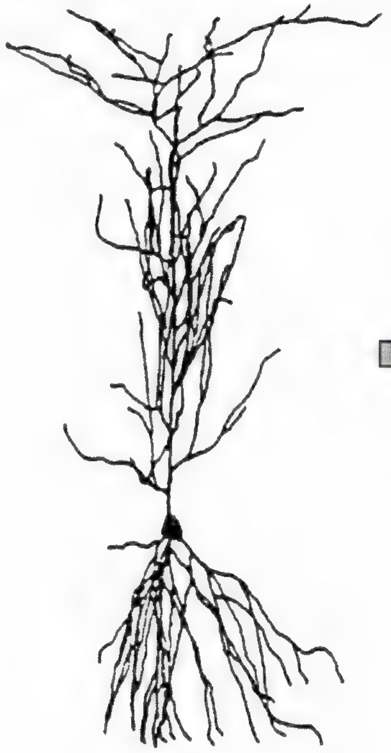
Kevin Archie
Yiota Poirazi

Technion, Haifa

Jackie Schiller
Alon Polsky

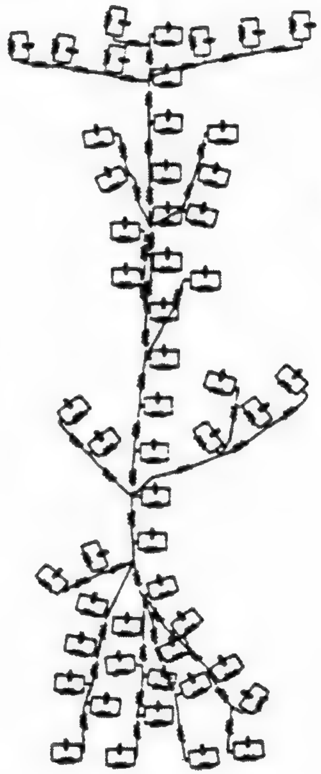
Modeling neurons as circuits

Neuron

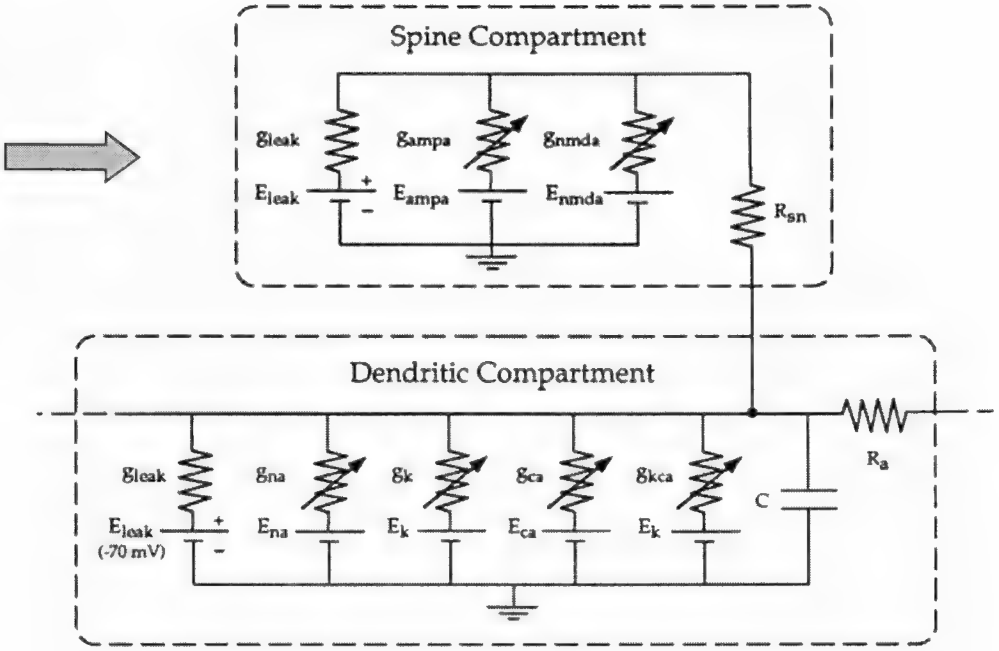


Zador et al. 1992

Circuit Model

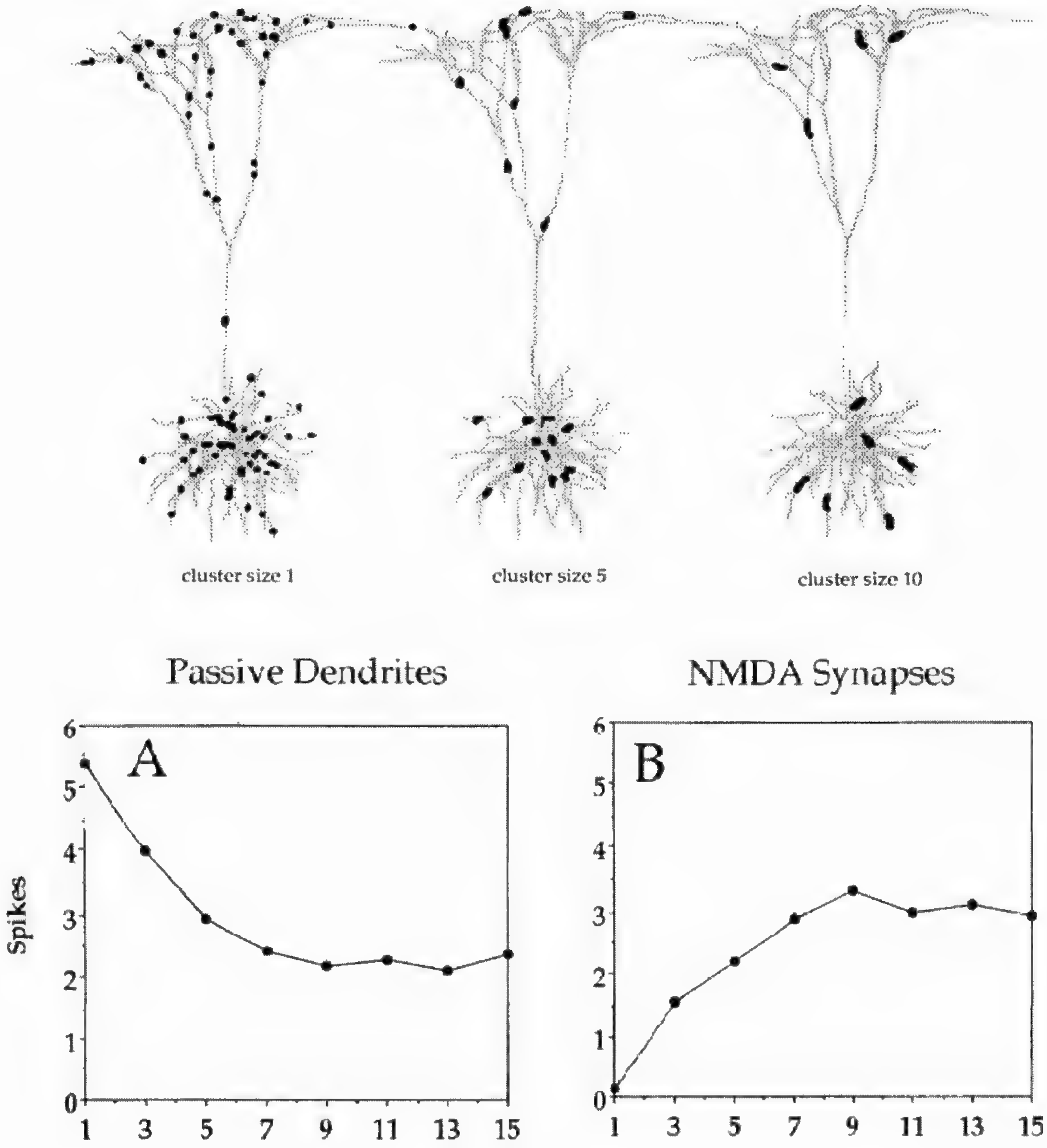


1 Dendrite + 1 Spine Compartment


Spine Compartment
g_{leak}
E_{leak}
+
-
g_{ampa}
E_{ampa}
g_{nmda}
E_{nmda}
R_{sn}
Dendritic Compartment
g_{leak}
E_{leak}
(-70 mV)
+
-
g_{na}
E_{na}
g_k
E_k
g_{ca}
E_{ca}
g_{kca}
E_k
C
R_a

We use NEURON
written by Mike Hines

“Cluster sensitivity”



Mel (1993)

The “clusteron”

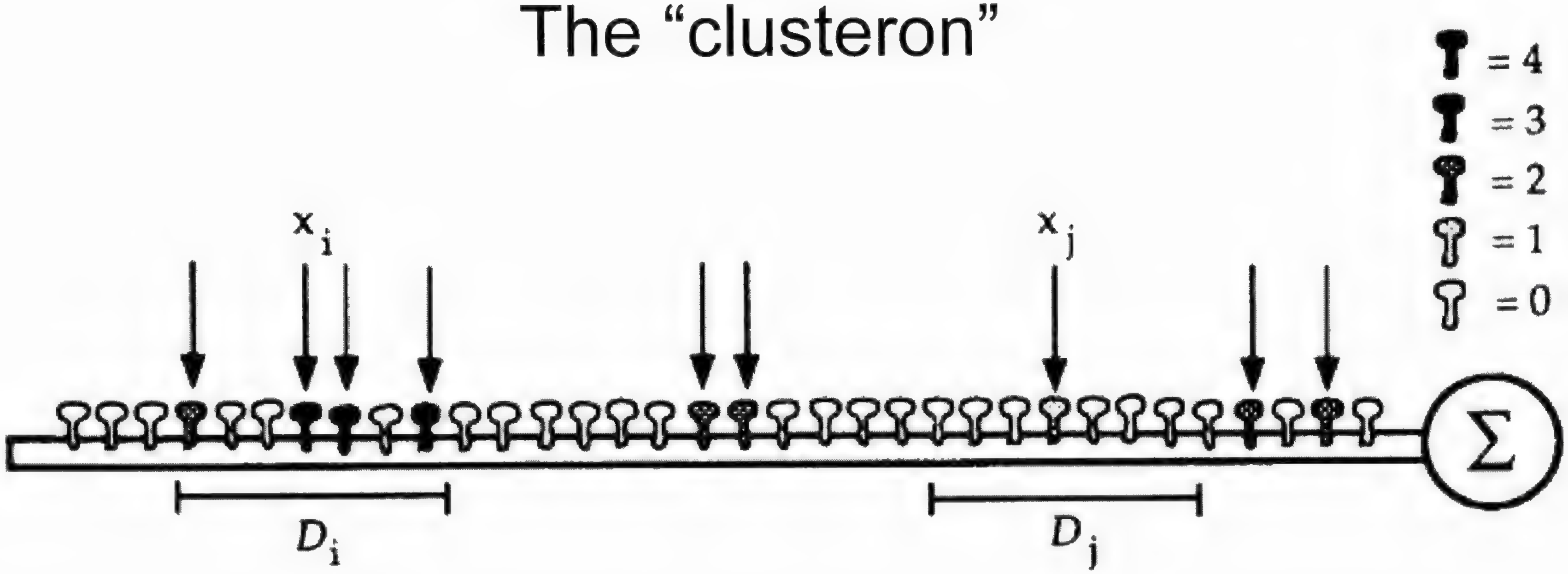


Figure 1: The *Clusteron*. Active inputs lines are designated by arrows; shading of synapses reflects synaptic activation a_i when $x_i \in \{0, 1\}$ and weights are set to 1.

$$a_i = w_i x_i \Big(\sum_{j \in \mathcal{D}_i} w_j x_j \Big)$$

Mel (1992)

Persistent Neural Activity and Neural Integrators: Experiments

Outline

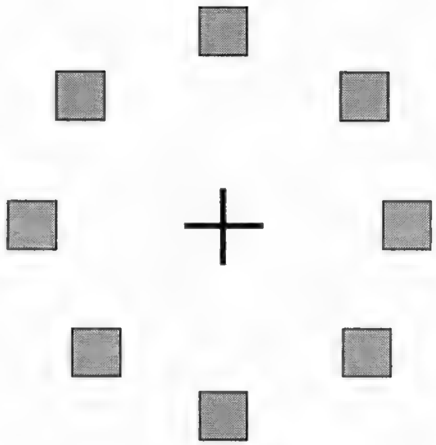
1. What is persistent neural activity? What is a neural integrator?
2. A model system: the Oculomotor Neural Integrator
3. Neural mechanisms of integration
    - Evidence for Network Mechanisms
    - Evidence for Cellular Mechanisms
4. Plasticity of the Integrator system

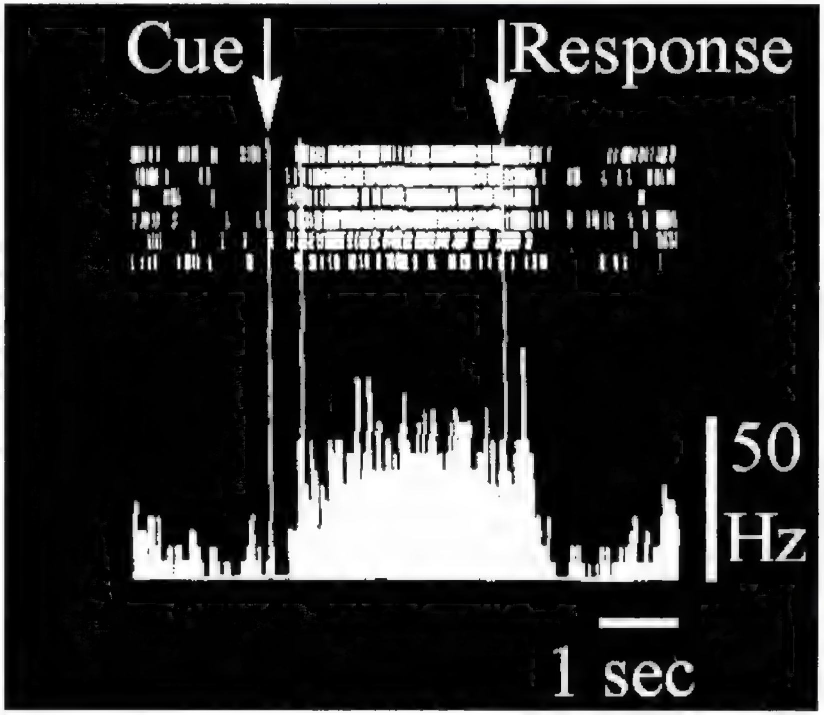
A memory problem

“Let us assume then that the persistence or repetition of a
reverberatory activity (or ‘trace’)
tends to induce lasting cellular changes that add to its stability.”

Hebb, 1948

Neural Correlate to Short-term Memory

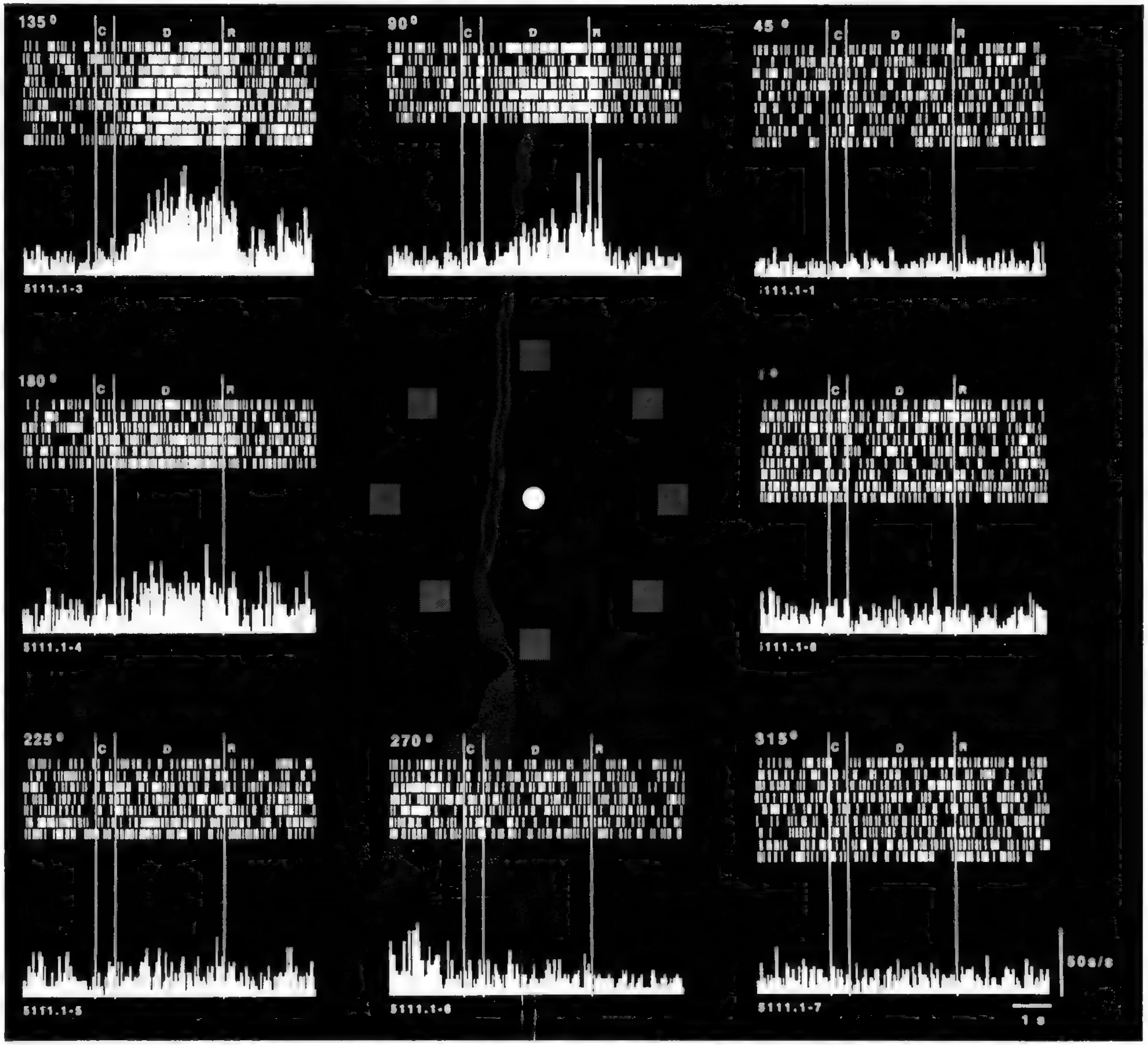




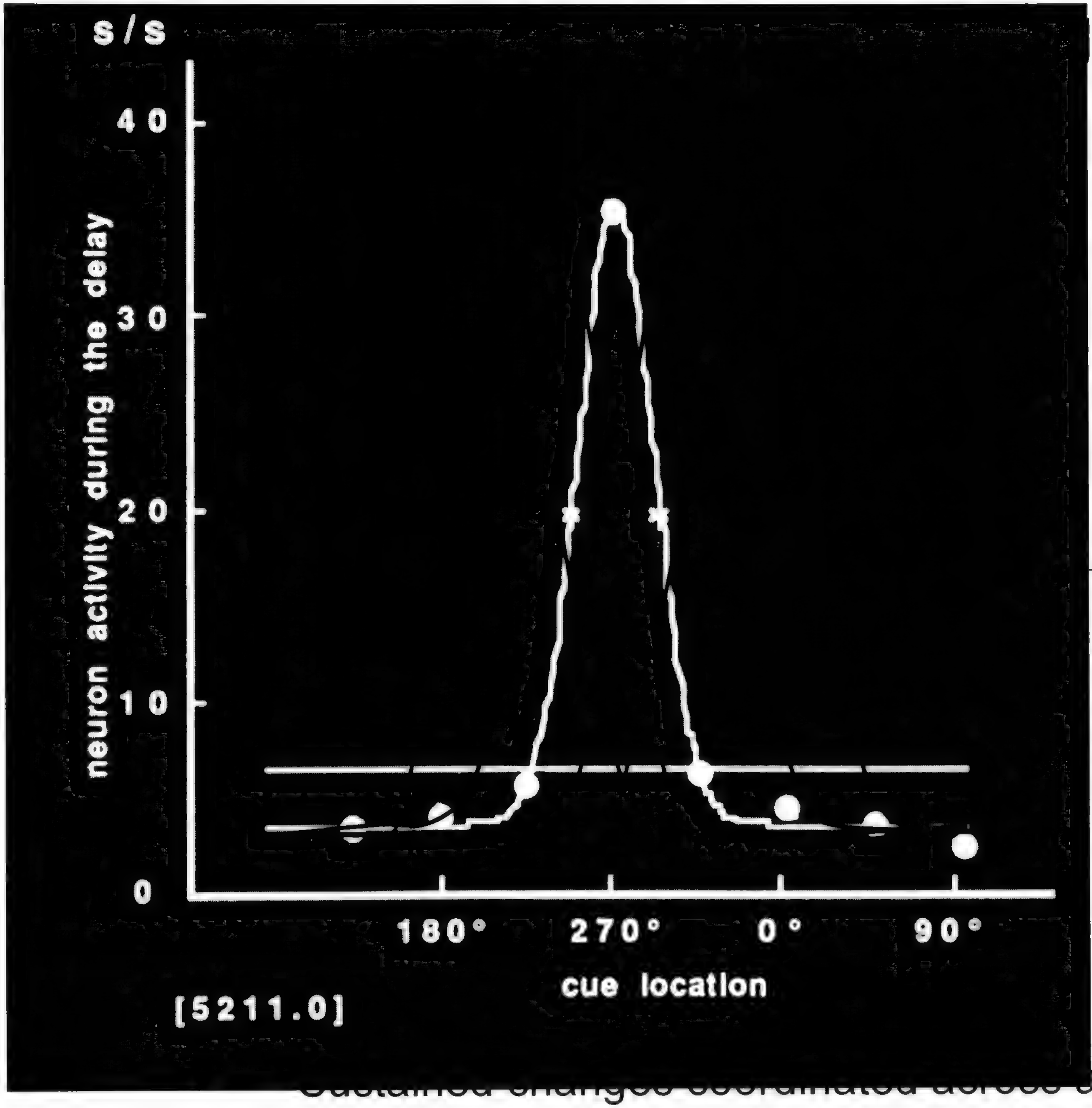
Fuster, Goldman-Rakic

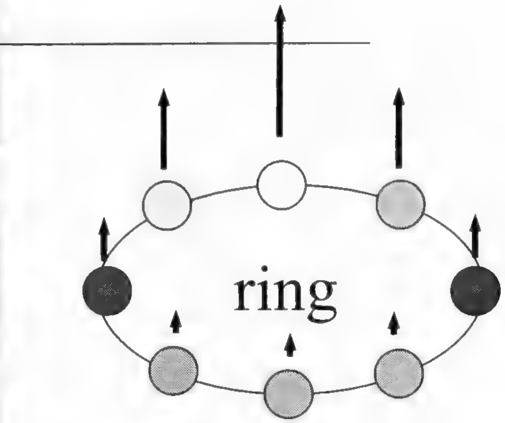
- Rapid changes between different rates
- Sustained changes in firing for many seconds

Directional Tuning of Memory Activity



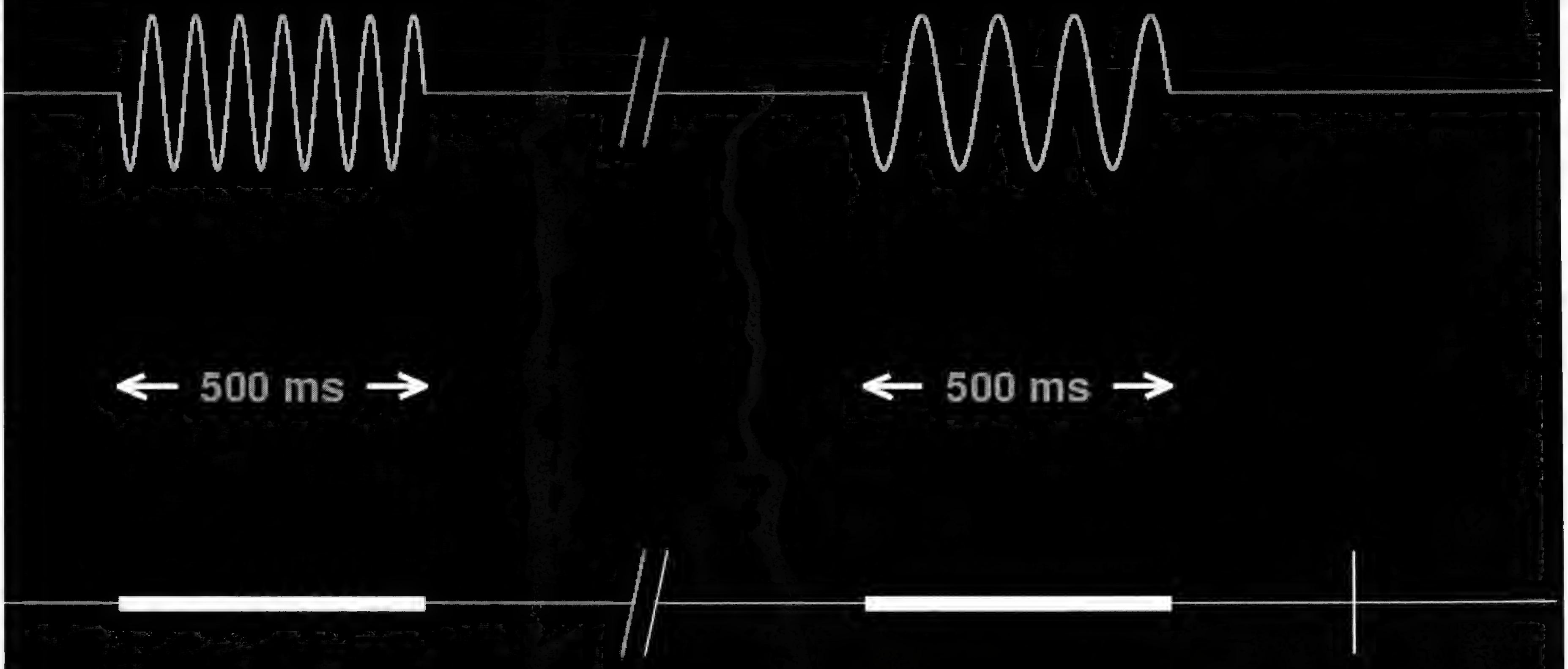
Memory





Sustained changes coordinated across a network

Mechanical flutter vibration (~10 to 50 Hz) memory task



← 500 ms →

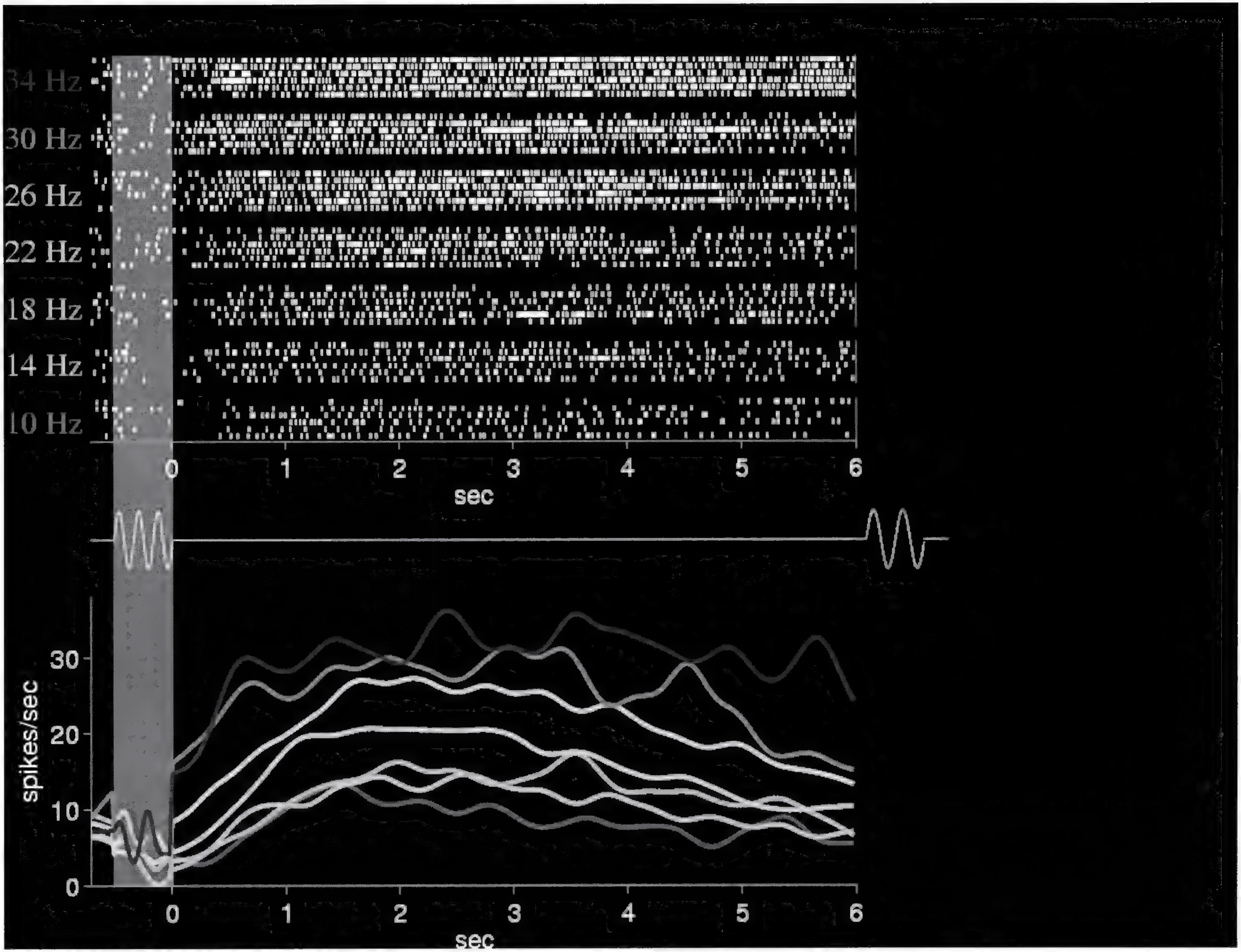
← 500 ms →

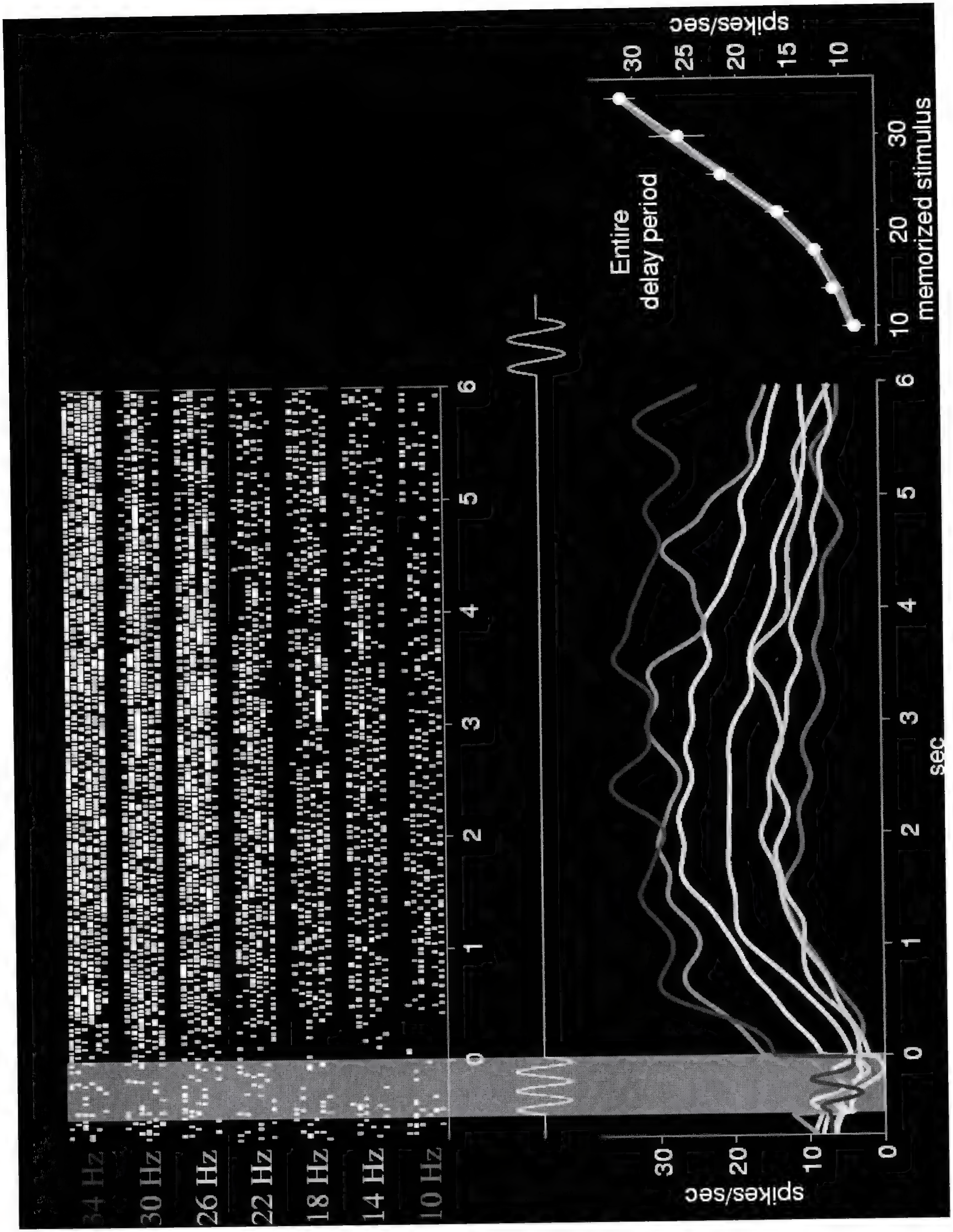
Memorized stim.
frequ. 1

Comparison stim.
frequ. 2

choose
pushbutton:
f 2 > f 1 ?

Romo, Brody, Hernandez, Lemus (1999), *Nature* **399**:470-473

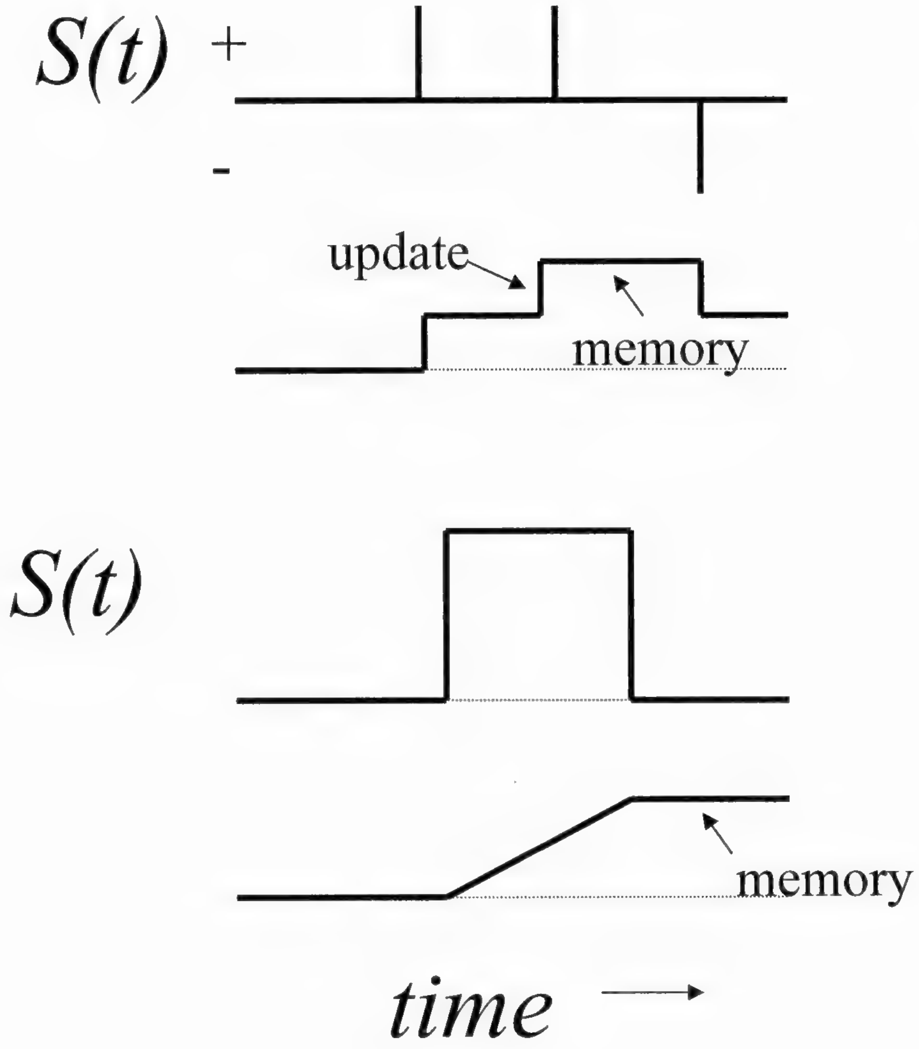

34 Hz
30 Hz
26 Hz
22 Hz
18 Hz
14 Hz
10 Hz
0
1
2
3
4
5
6
sec
spikes/sec
30
20
10
0
0
1
2
3
4
5
6
sec


34 Hz
30 Hz
26 Hz
22 Hz
18 Hz
14 Hz
10 Hz
0
1
2
3
4
5
6
sec
spikes/sec
0
10
20
30
Entire
delay period
memorized stimulus
10
20
30
spikes/sec
10
15
20
25
30

From Memory to Integration

- In short-term memory systems, stored patterns represent input variables
- In neural integrators, stored patterns represent accumulated input

Idealized Responses


S(t)
+
-
update
memory
S(t)
memory
time

pulse

⇓

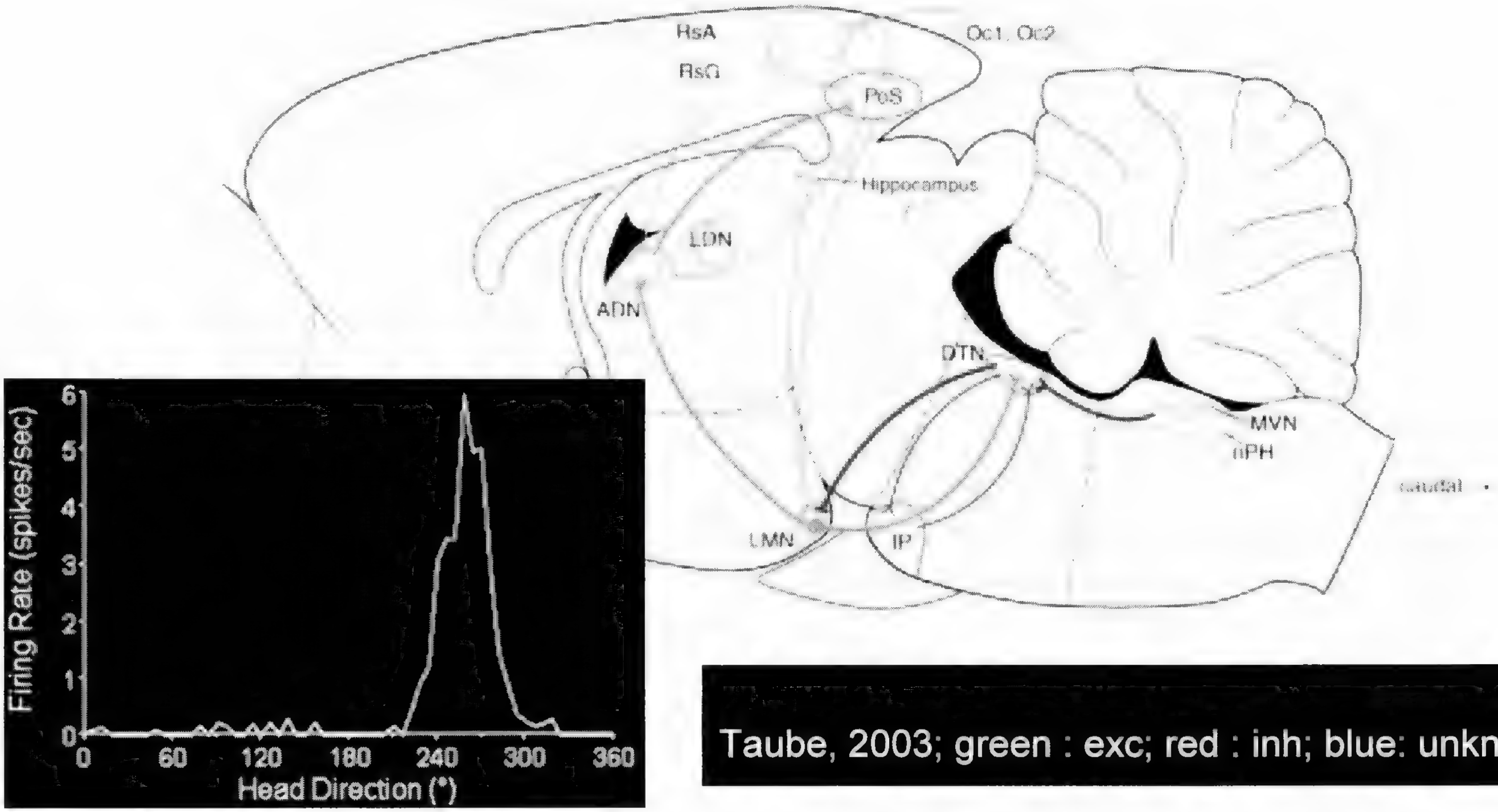
step

constant

⇓

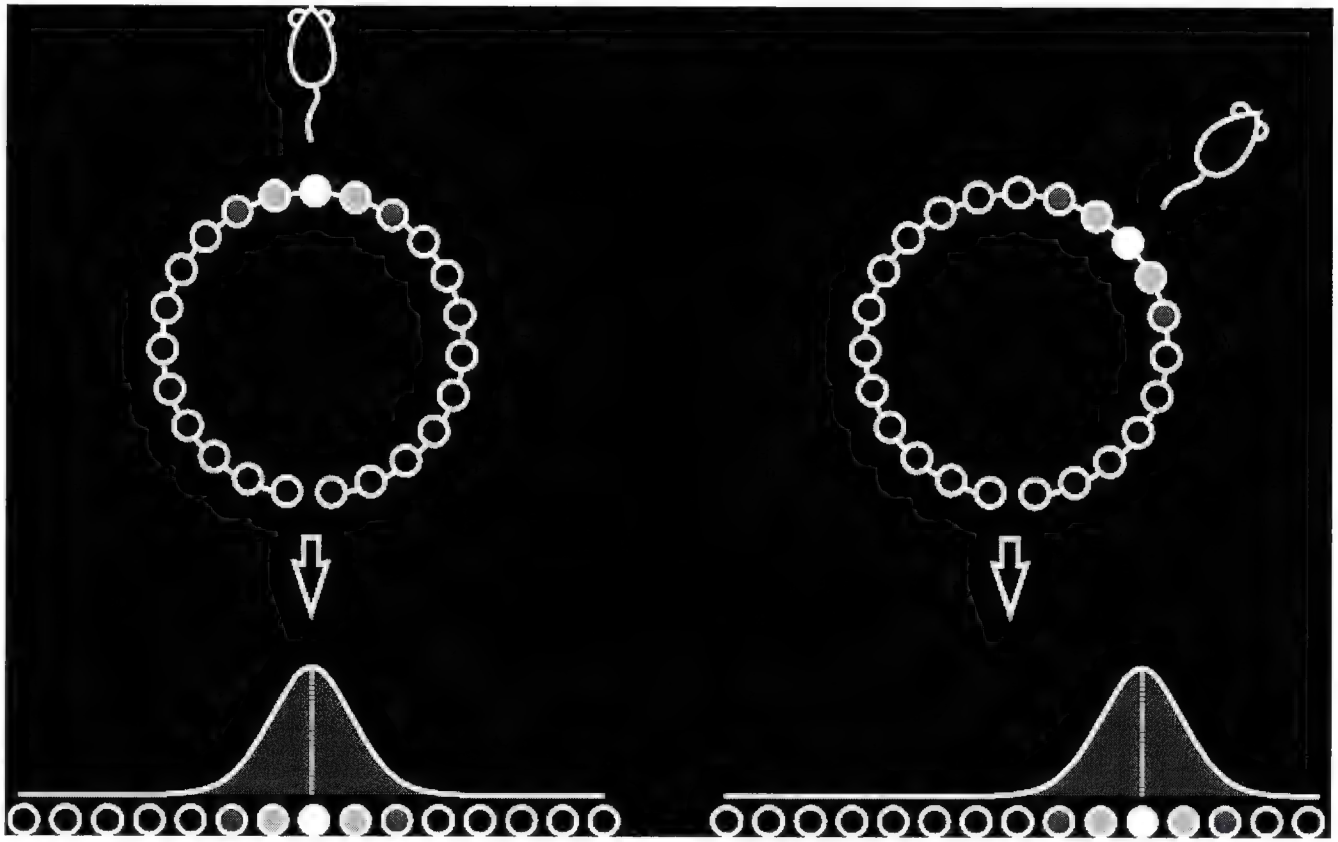
ramp

The Head Direction System


RsA
RsG
Oc1, Oc2
PoS
Hippocampus
LDN
ADN
DTN
MVN
NPH
caudal
LMN
IP
Firing Rate (spikes/sec)
0
1
2
3
4
5
6
0
60
120
180
240
300
360
Head Direction (°)

Taube, 2003; green : exc; red : inh; blue: unknown

Ring Structure in Population Activity

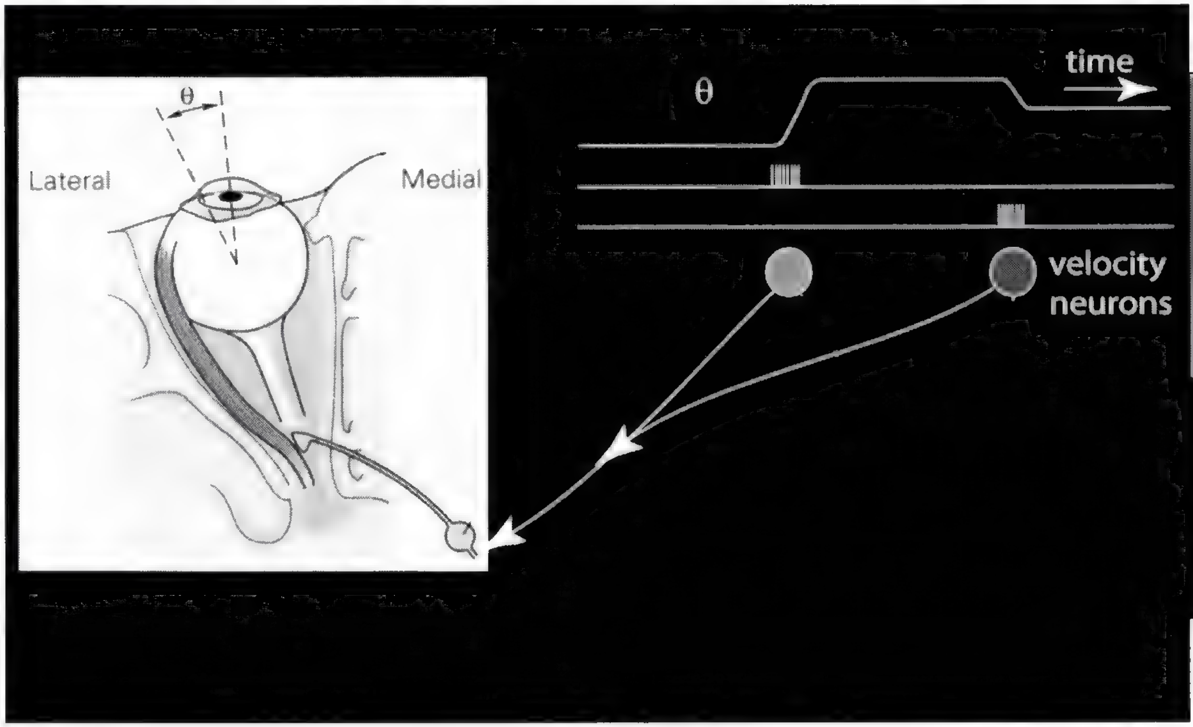


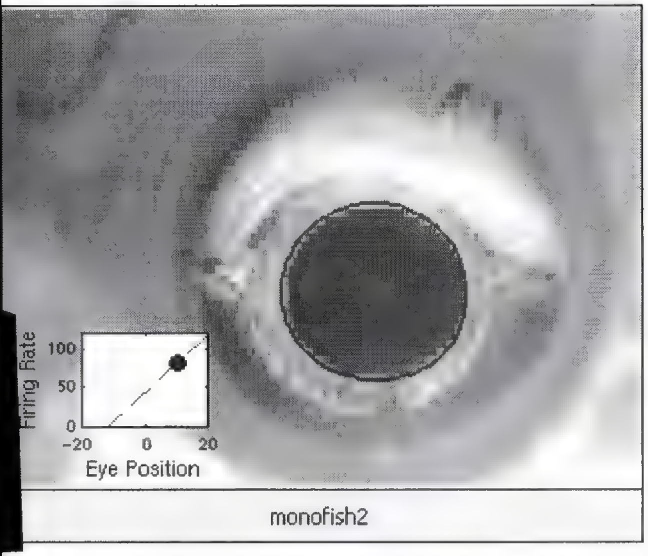
$$\int \omega \, dt = \theta$$

angular head velocity

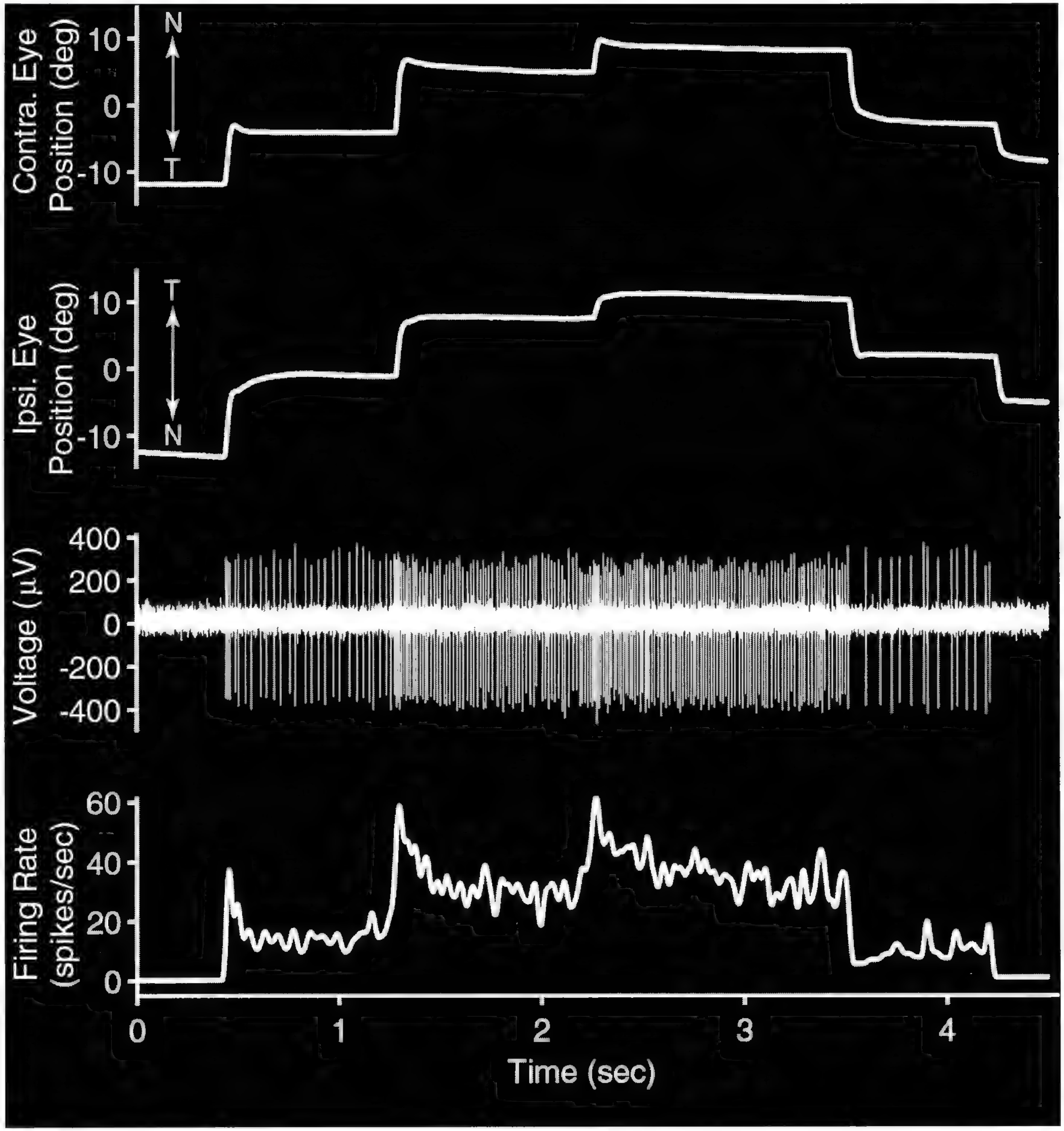
angular head position

Controlling Eye Position


θ
Lateral
Medial
θ
time
velocity
neurons


Firing Rate
100
50
0
-20
0
20
Eye Position
monofish2

- Sensory signals and initial commands are encoded as velocities
- Specialized cells must convert this velocity information to position commands



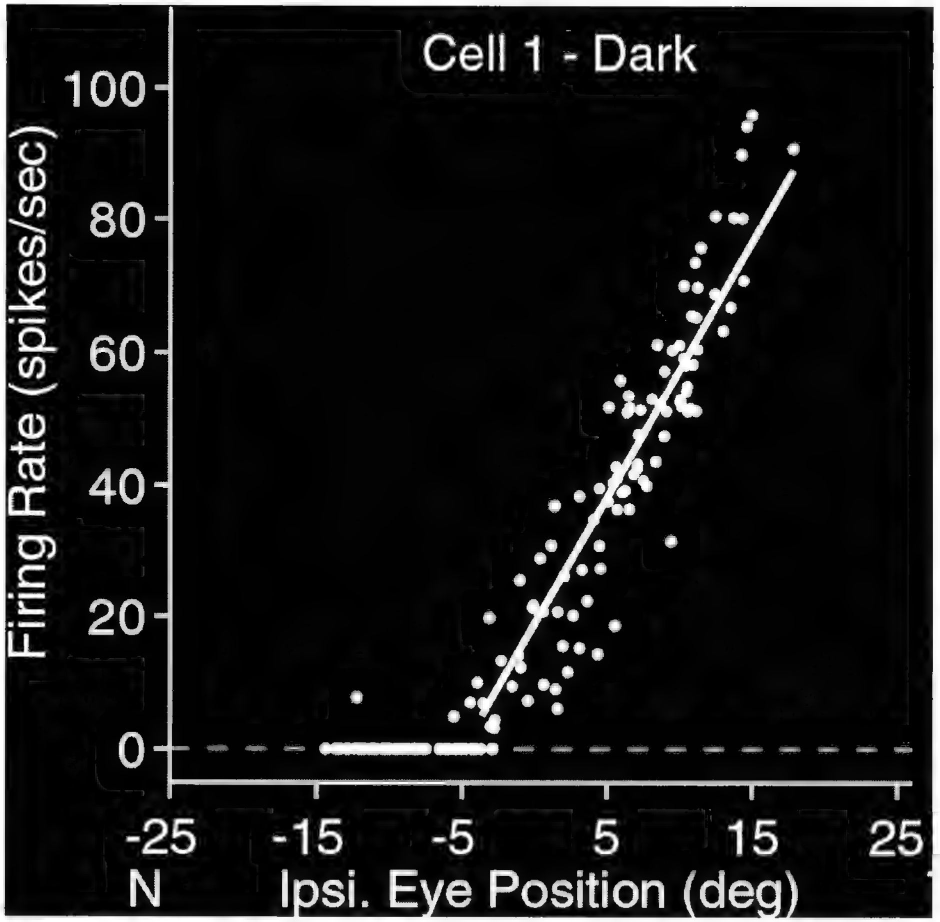
activity

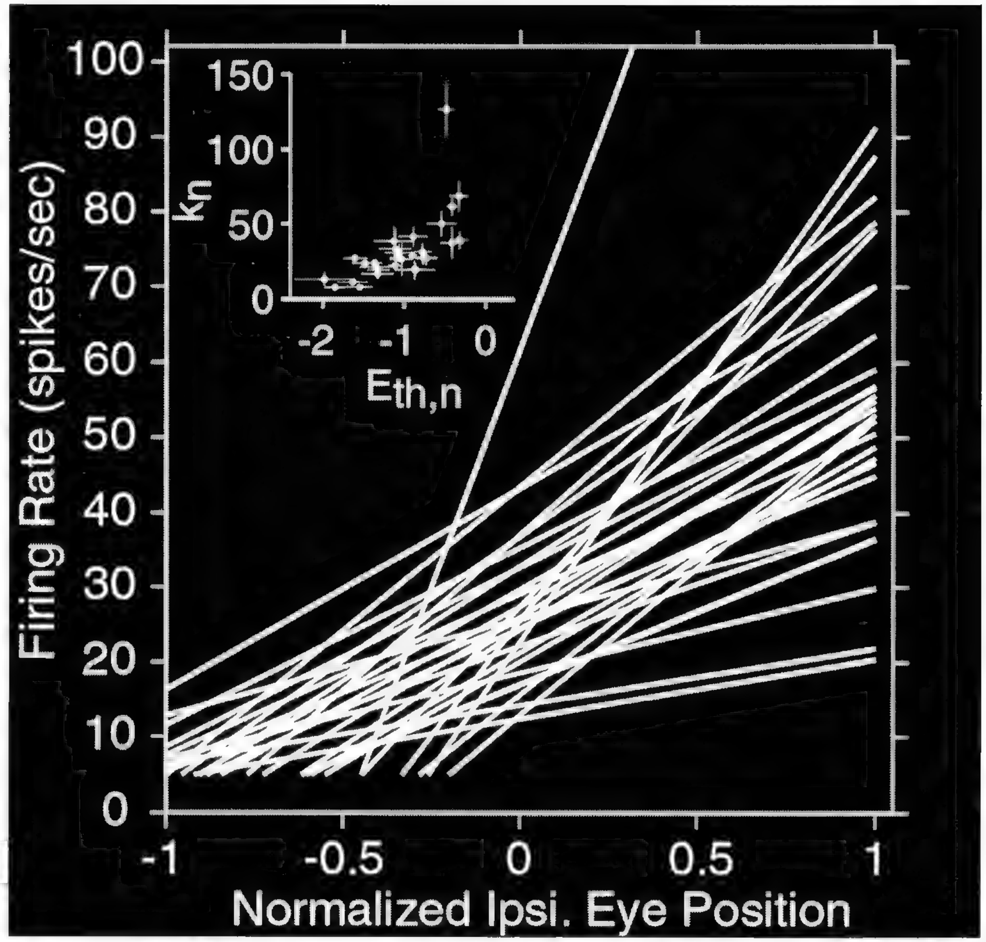
eye position)

On/off

burst

input





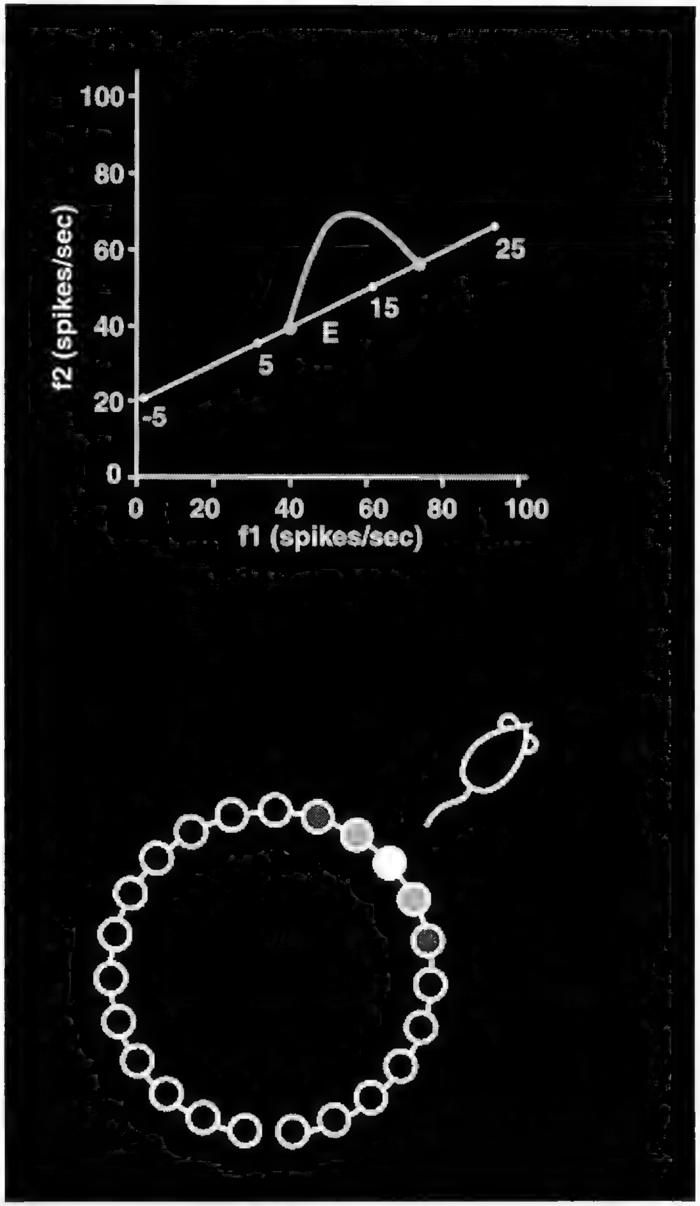
firing rate ↘

threshold ↙

monotonic
encoding

↗
position sensitivity

Coding Summary

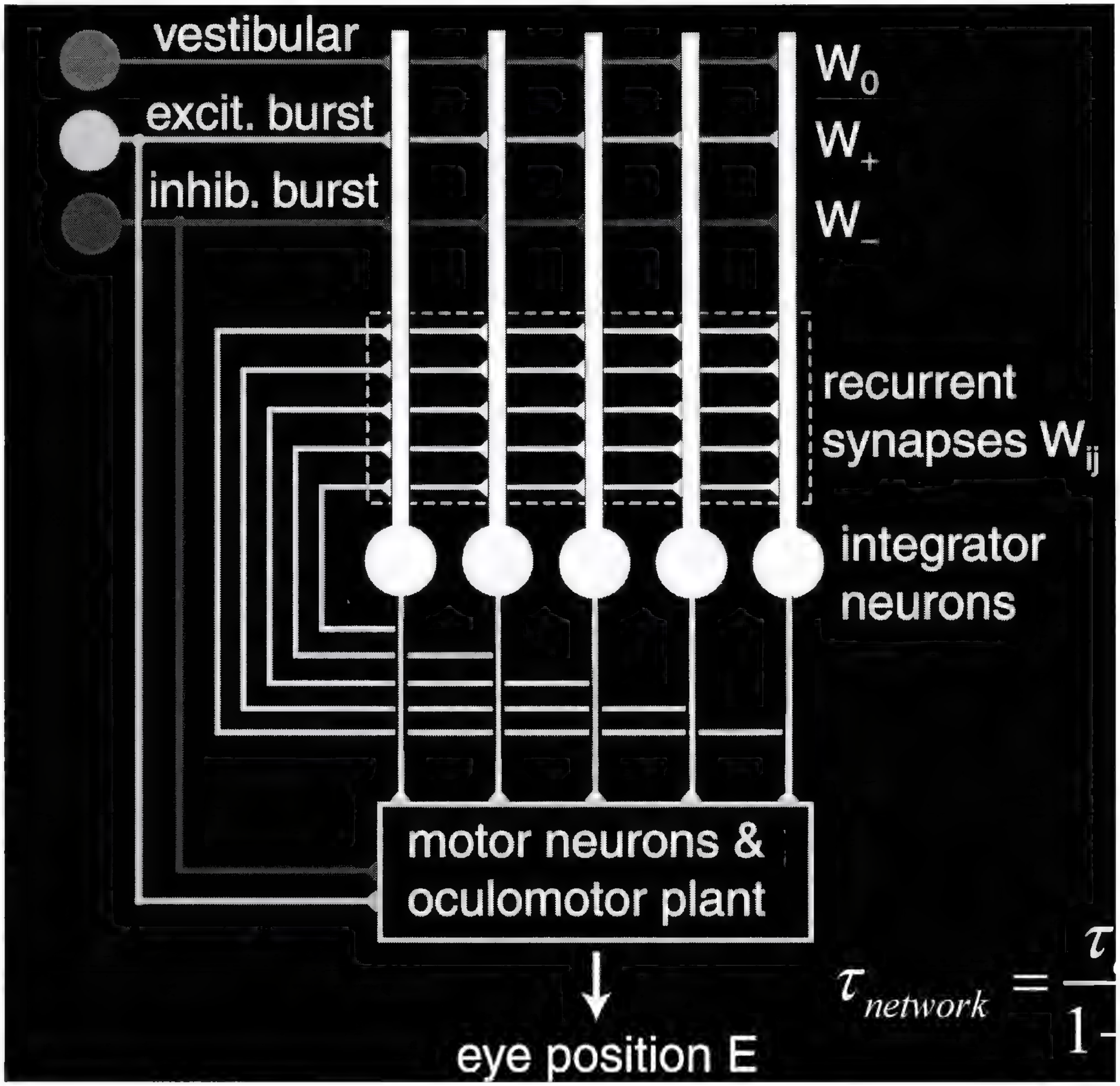

f2 (spikes/sec)
100
80
60
40
20
0
f1 (spikes/sec)
0
20
40
60
80
100
-5
5
E
15
25

- Stored representation of either a stimulus parameter, or integral of a stimulus parameter
- Sustained elevated (or reduced) firing rates lasting up to tens of seconds (persistence time ~ 10 sec) long outlasting stimulus
- Rapid onset, offset
- Mean firing rate an analog variable
- Stable *pattern* of activity (population code)

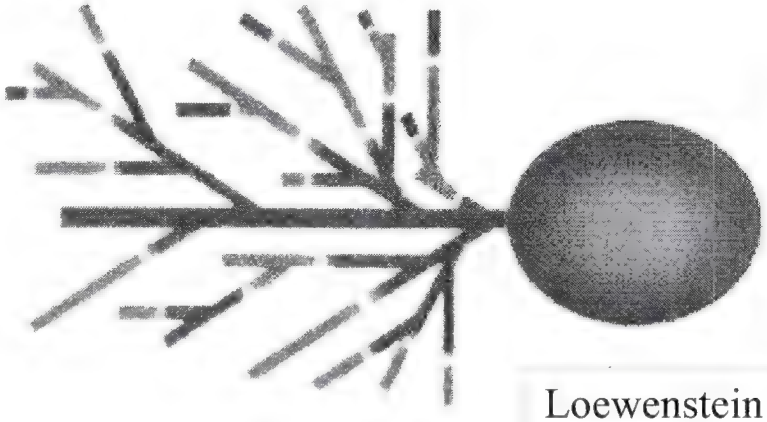
Mechanisms of Neural Integration

The Network Positive Feedback Hypothesis

- Rosen (1972)
- Kamath and Keller (1976)
- Cannon, Robinson Shamma (1983)
- Cannon, Robinson (1987)
- Arnold, Robinson (1991)
- Seung (1996)
- Seung, Lee, Reis, Tank (2000)
- Sklavos and Moschovakis (2002)


vestibular
excit. burst
inhib. burst
W_0
W_+
W_-
recurrent synapses W_{ij}
integrator neurons
motor neurons & oculomotor plant
eye position E
$\tau_{network} = \frac{\tau}{1}$

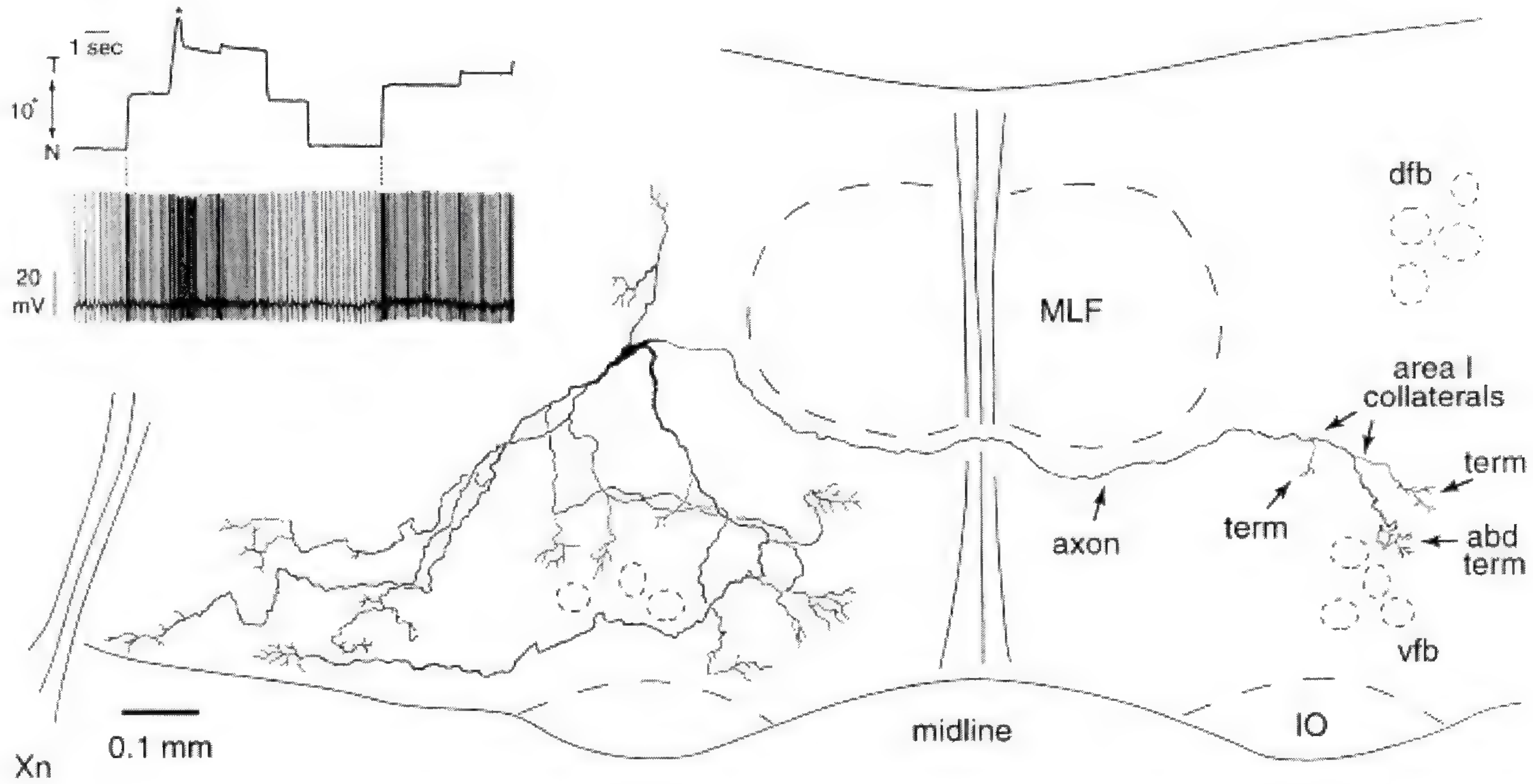
Potential Cellular Mechanisms



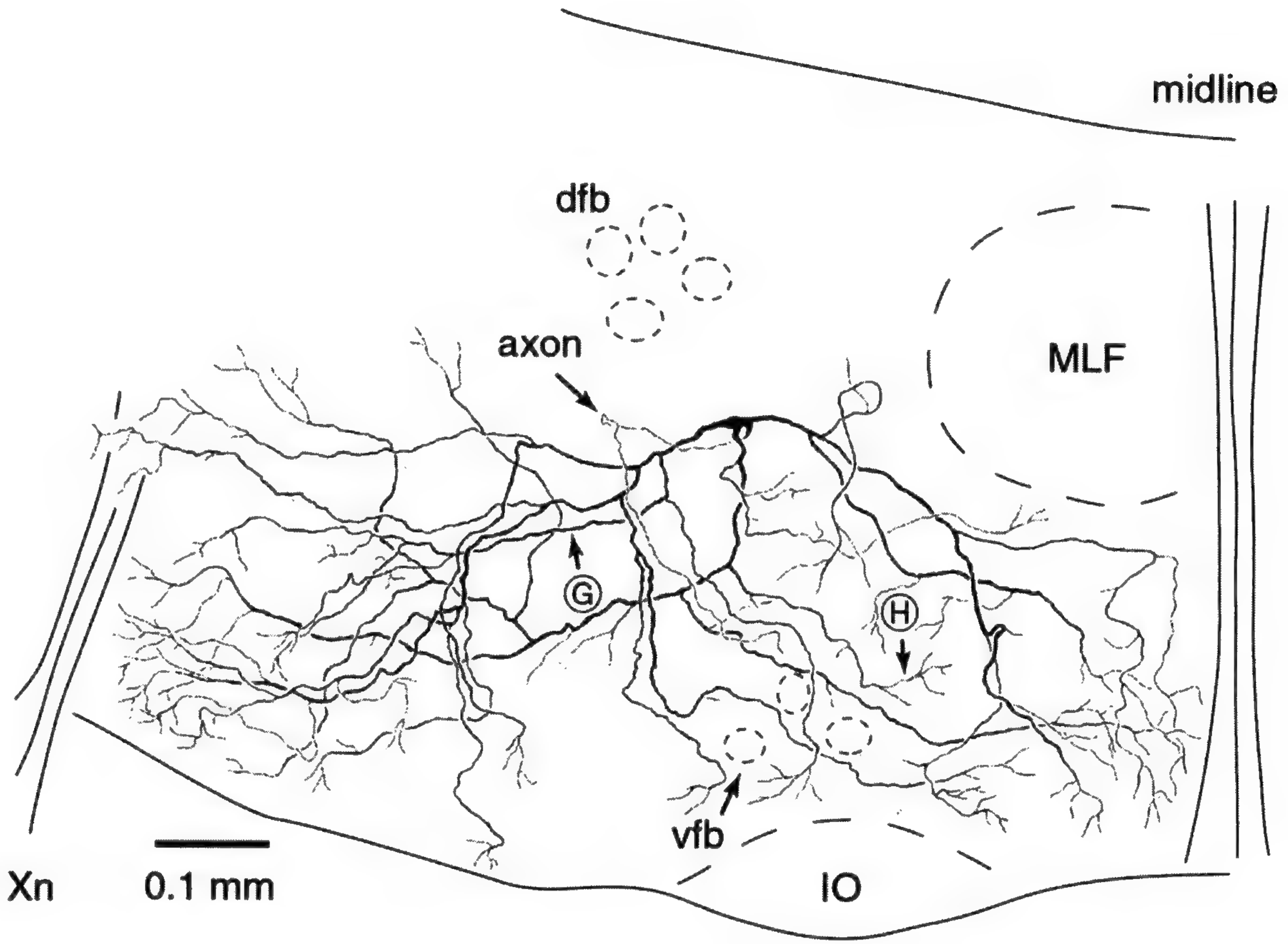
Loewenstein and
Sompolinsky, 2003

1. Intrinsic oscillators
2. Threshold regulation
3. Plateau potentials
4. Traveling calcium fronts

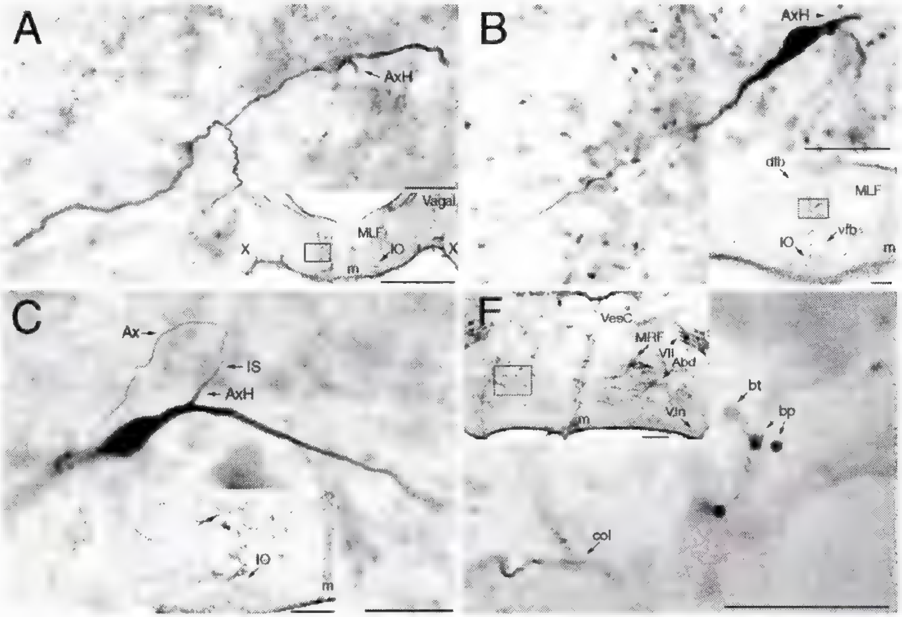
Cells with Opposite-side Projections


1 sec
T
10°
N
20
mV
dfb
MLF
area I
collaterals
term
term
abd
term
axon
vfb
Xn
0.1 mm
midline
IO

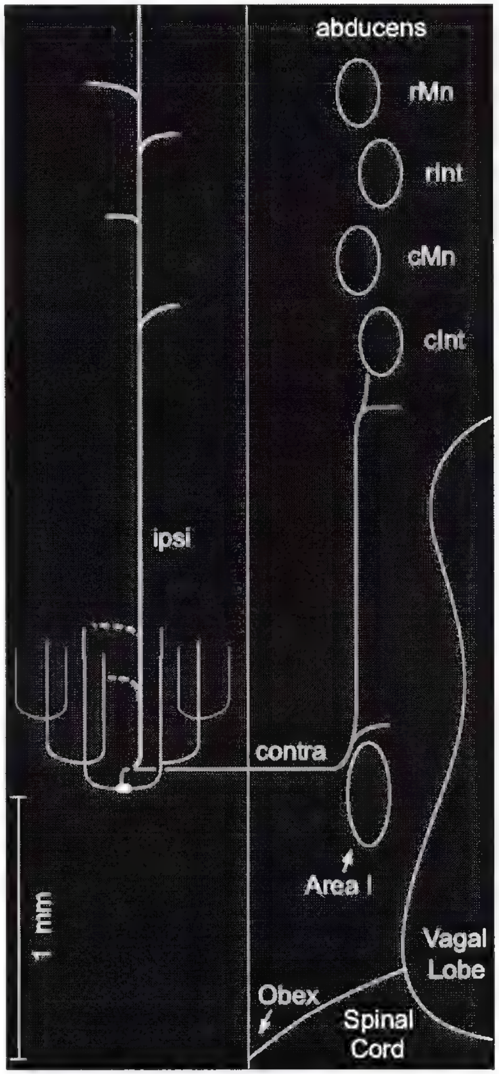
Cells with Same-side Projections


midline
dfb
axon
MLF
G
H
vfb
Xn
0.1 mm
IO

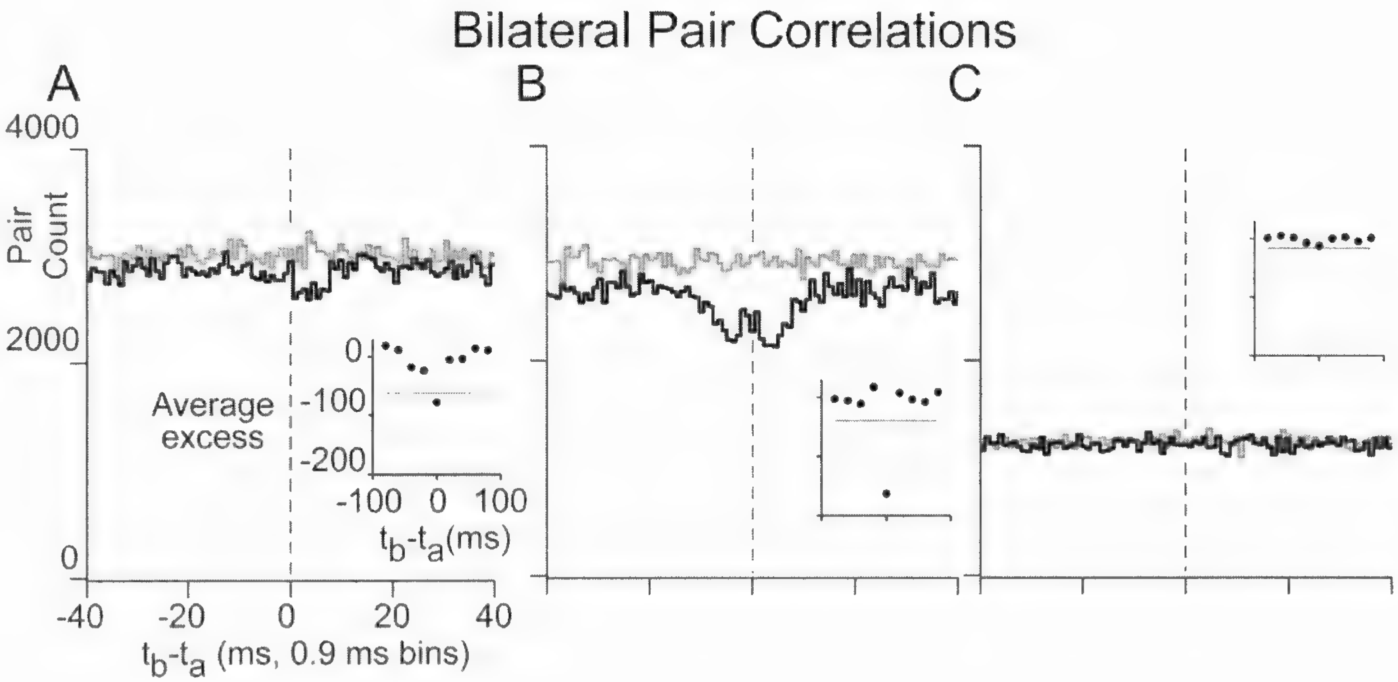
The Anatomy of an Integrator


A
AxH
Vagal
MLF
IO
X
X
m
B
AxH
dfb
MLF
vfb
IO
m
C
Ax
IS
AxH
IO
m
F
VesC
MRF
VII
Abd
VIn
m
bt
bp
col

1. Collaterals on crossing projection
2. Preliminary evidence for local collaterals on same side.
3. Projections to the motoneurons

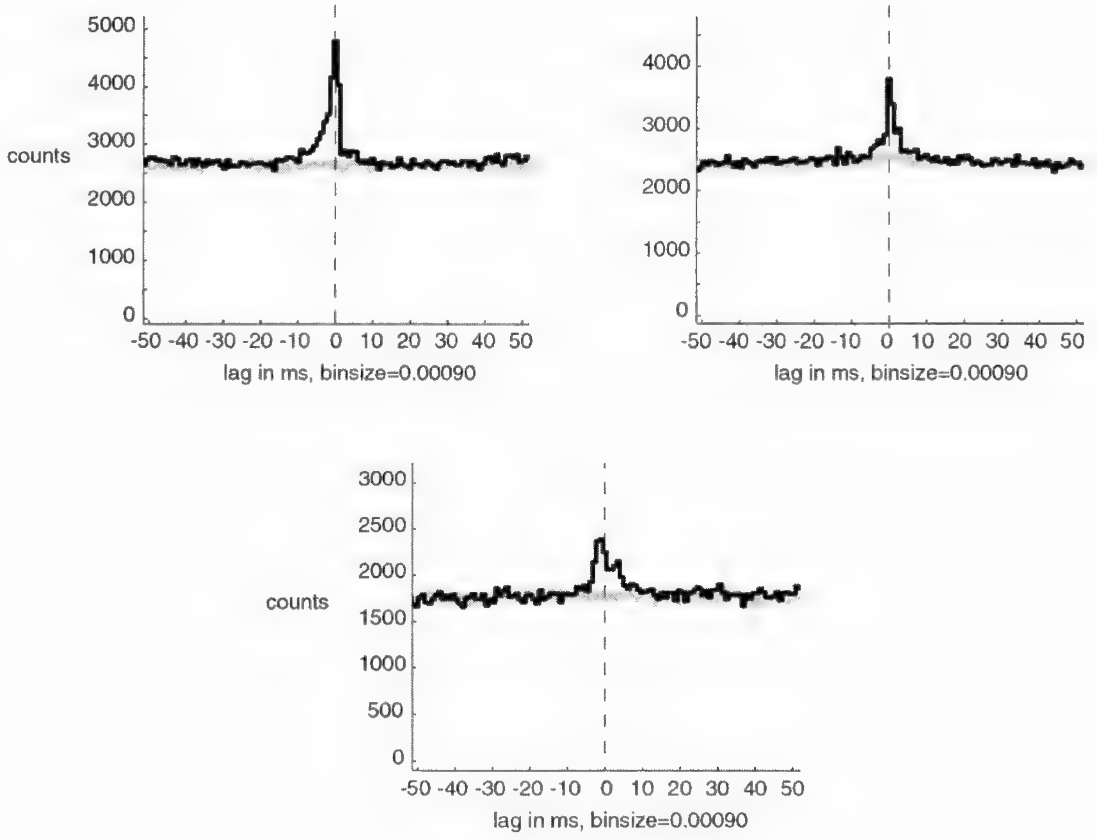

abducens
rMn
rInt
cMn
cInt
ipsi
contra
Area I
Vagal Lobe
Obex
Spinal Cord
1 mm

Opposite-side correlations



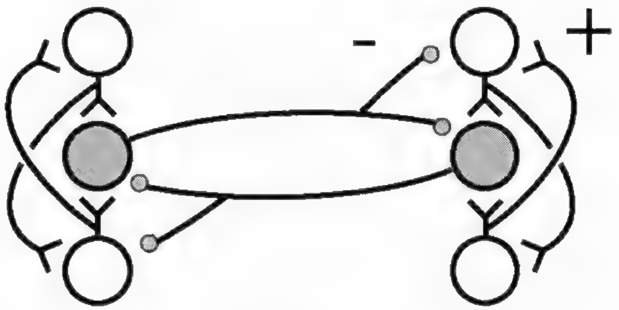
Correlations agree with opposite-side inhibitory connections

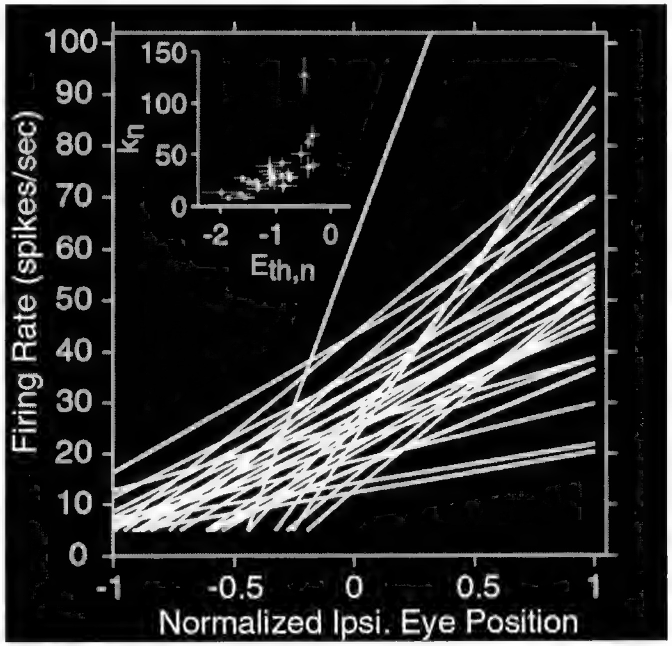
Same-side correlations



Correlations agree with same-side excitation

A Network for Integration

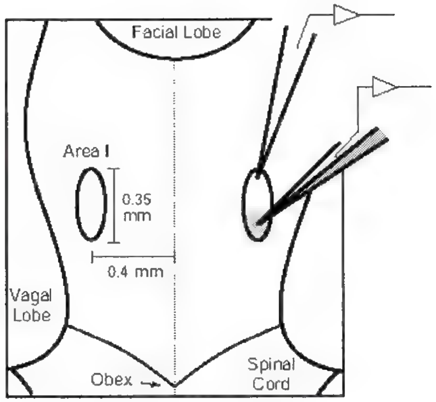


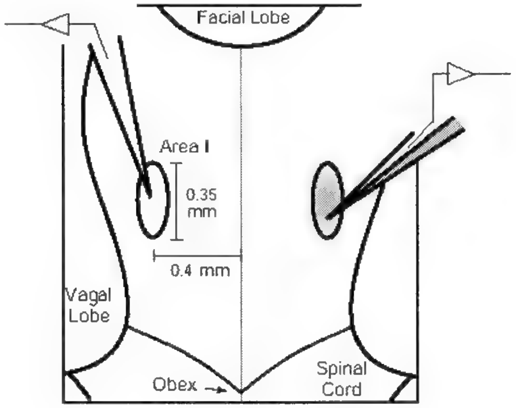


- Excitatory paths on same side, Inhibitory paths for opposite side
- Lateralization of function, so excitatory between on/on and inhibitory between on/off
- Asymmetric distribution of position thresholds leads to nonlinear population activity profile

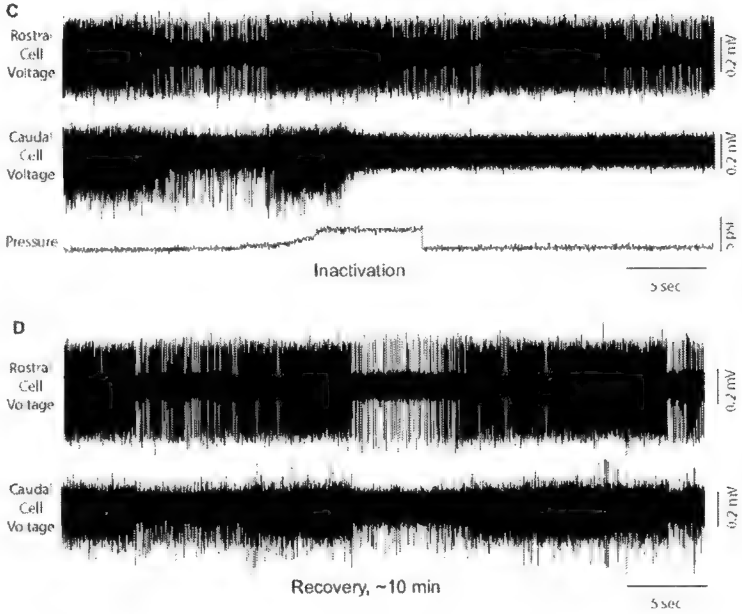
Functional Dissection of Circuitry

- Test same-side excitation by silencing caudal cells (dendritic tree sweeps rostrally)


Facial Lobe
Area I
0.35 mm
0.4 mm
Vagal Lobe
Obex
Spinal Cord

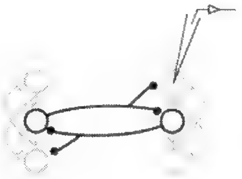

Facial Lobe
Area I
0.35 mm
0.4 mm
Vagal Lobe
Obex
Spinal Cord

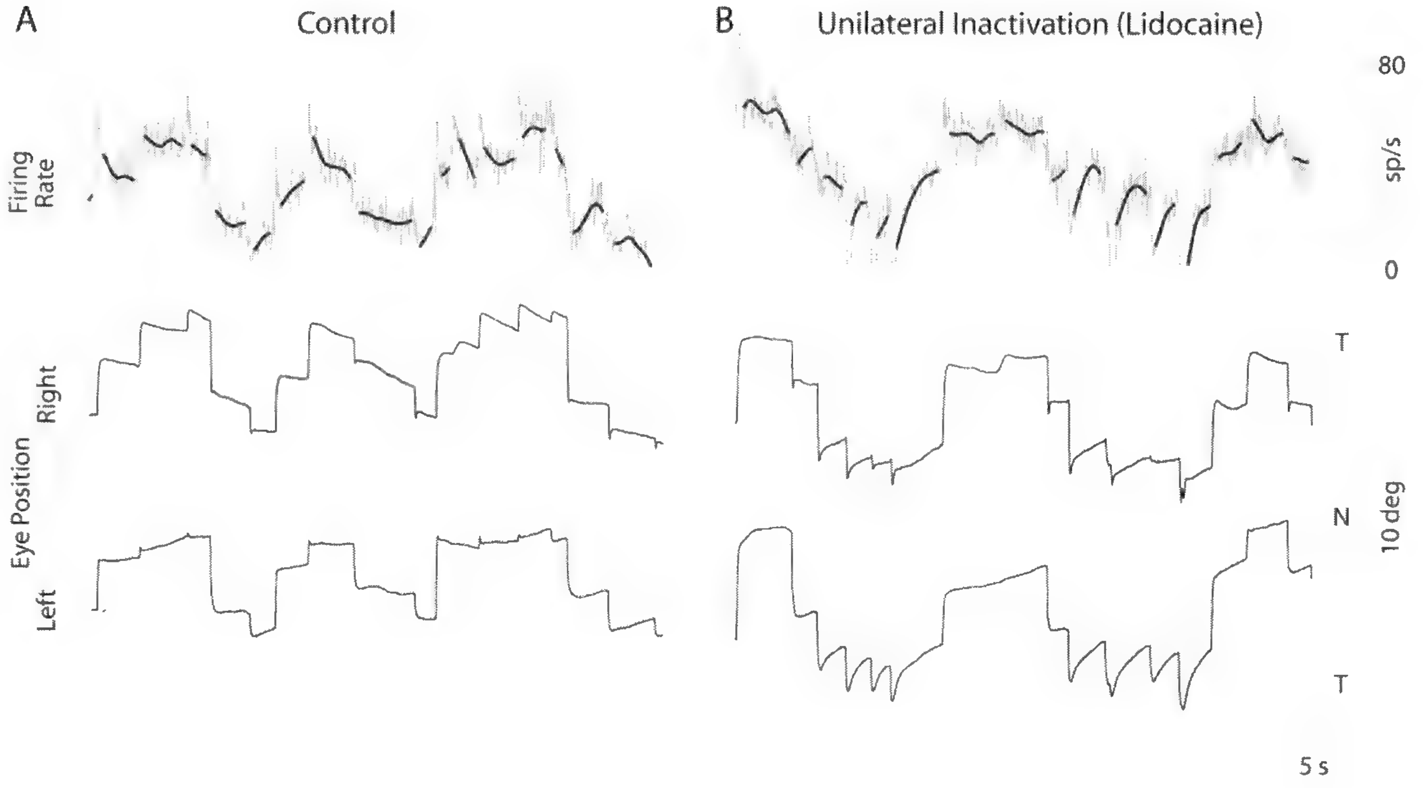
- Test crossing inhibition by silencing contralateral cells



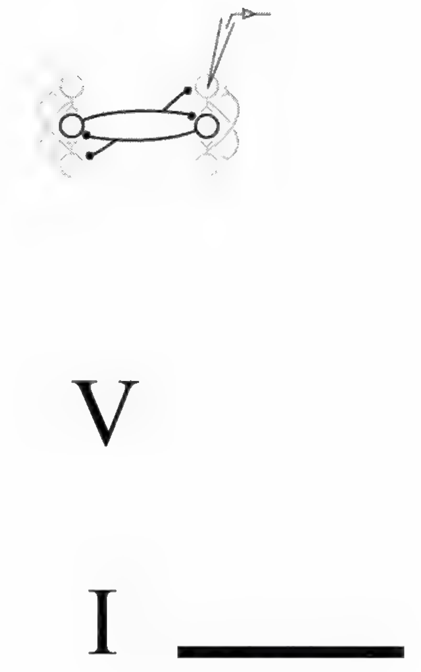
- An example of network perturbation:
- Recovery in 5-10 minutes
- Estimated ~1/3 silenced

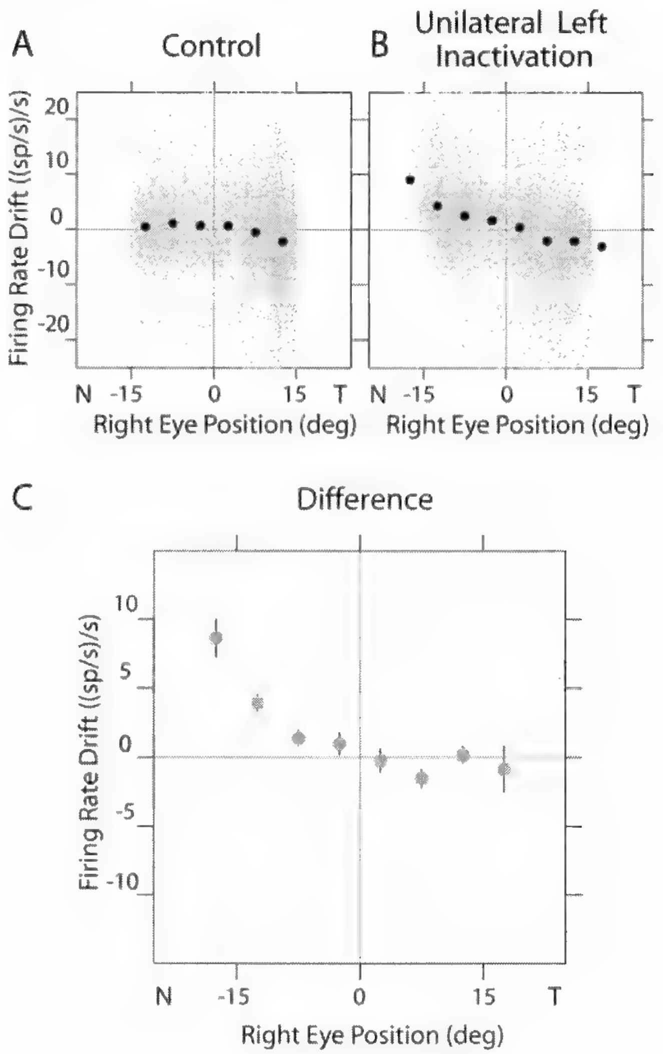
Role of Crossing Paths





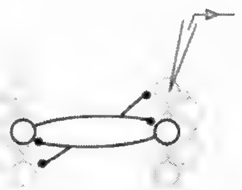
Role of Crossing Paths

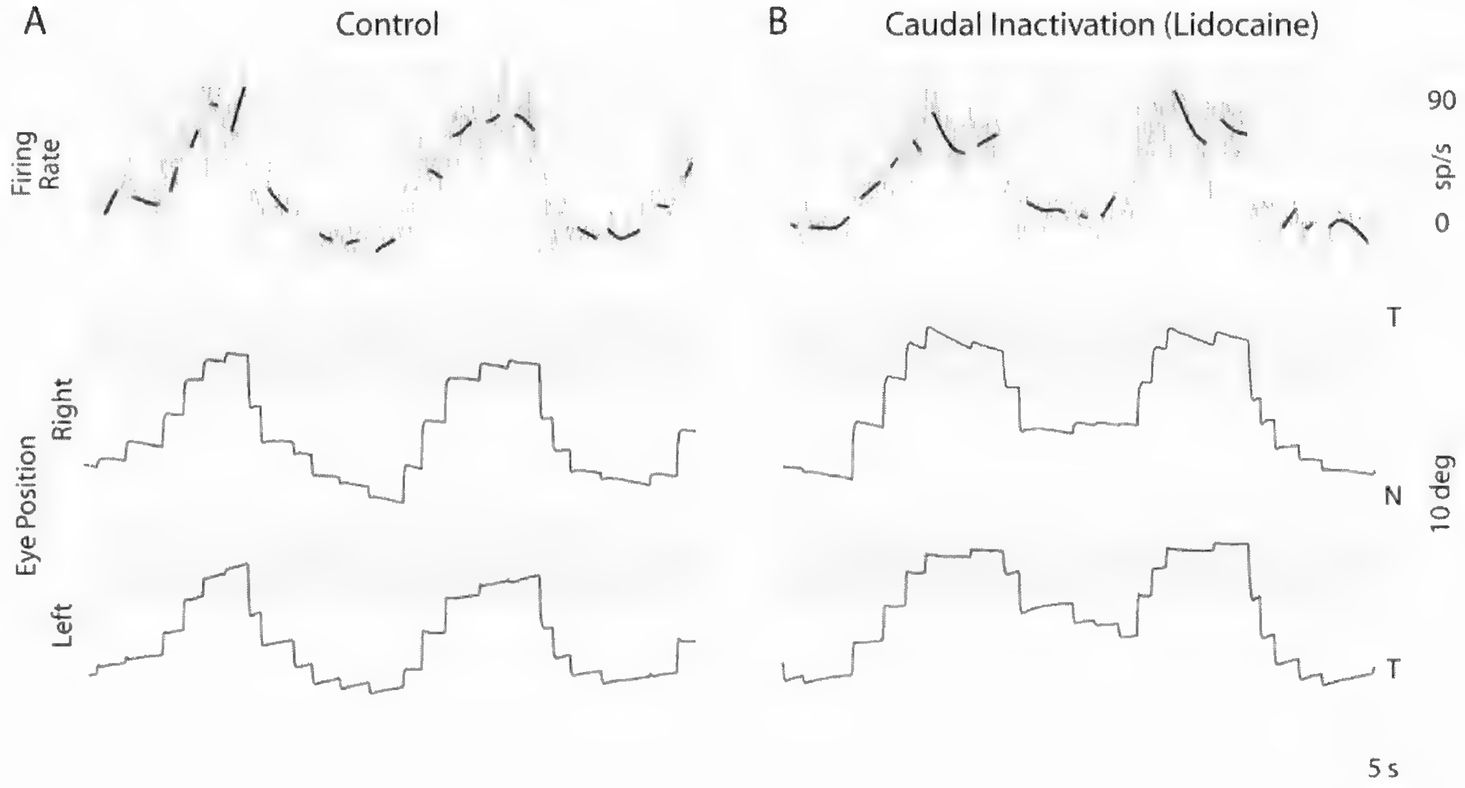




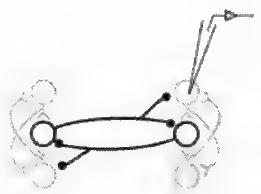
Memory decay, but only when Below an equilibrium position

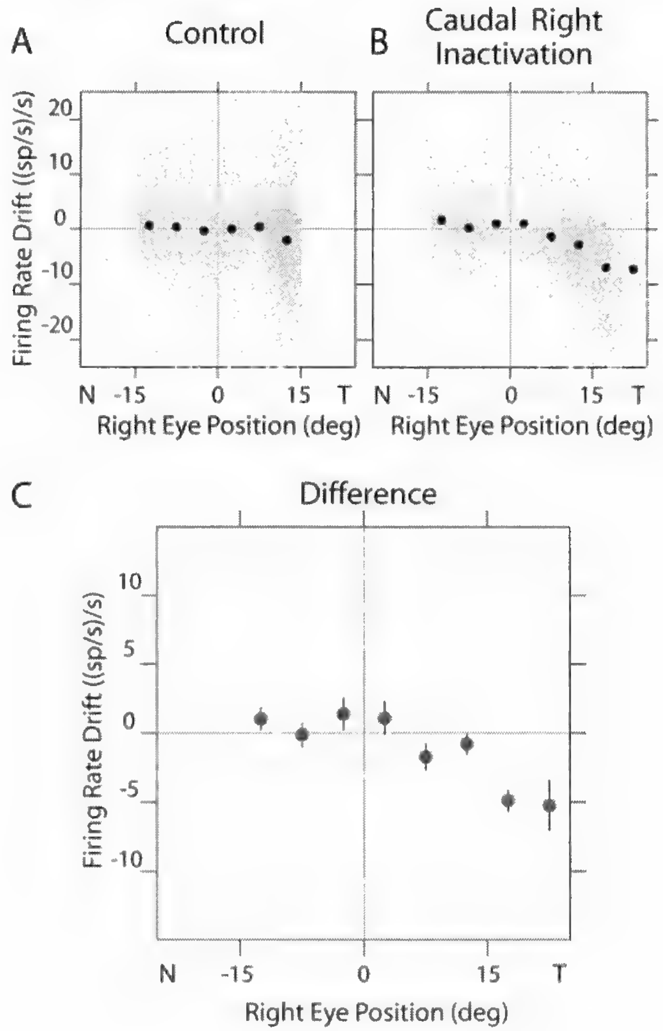
Role of Same-side Paths





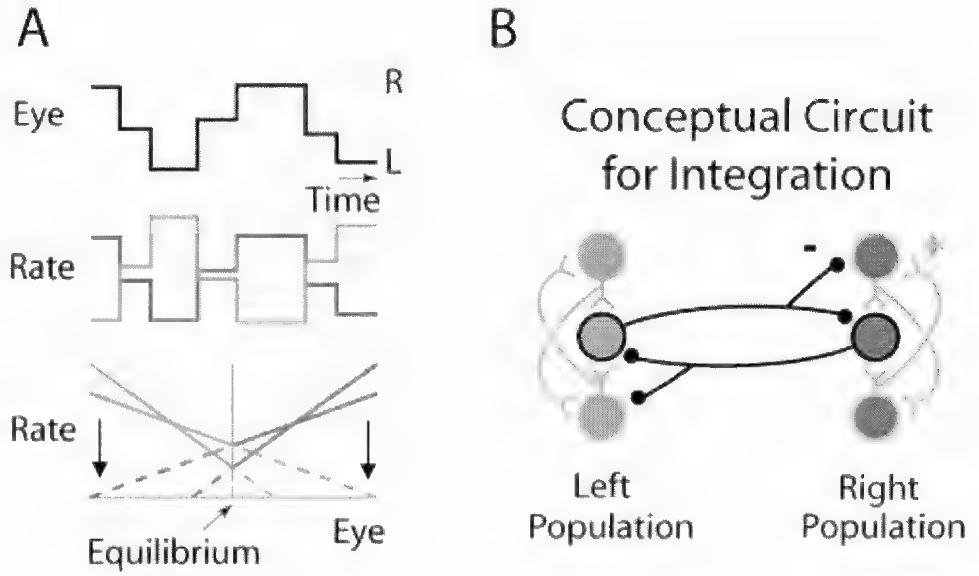
Role of Same-side Paths





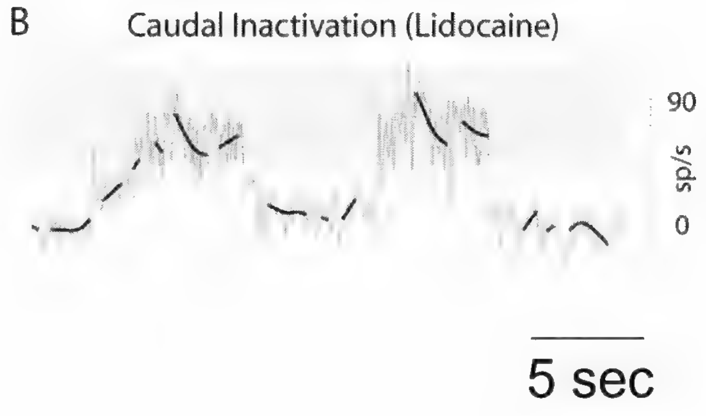
Memory decay, but only when Above an equilibrium position

Excitation for Persistance, Inhibition for Coupling


A
Eye
R
L
Time
Rate
Rate
Equilibrium
Eye
B
Conceptual Circuit for Integration
-
+
Left Population
Right Population

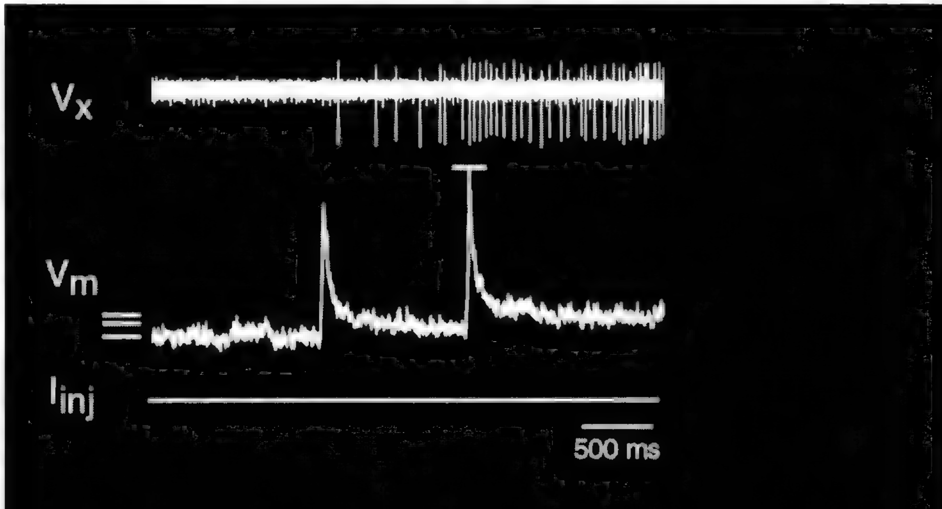
Persistance alternately generated by one population then the other as the eye moves from one side of equilibrium to the other. Inhibition acts in feedforward manner, keeps the two sides in register.

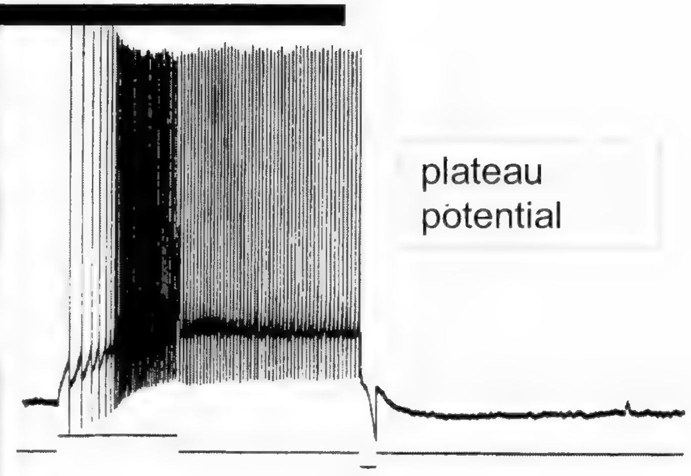
The Robustness Puzzle



- Only part of the dynamic range becomes leaky
- Time constants an order of magnitude longer than expected

Does the saccadic input trigger a plateau potential?





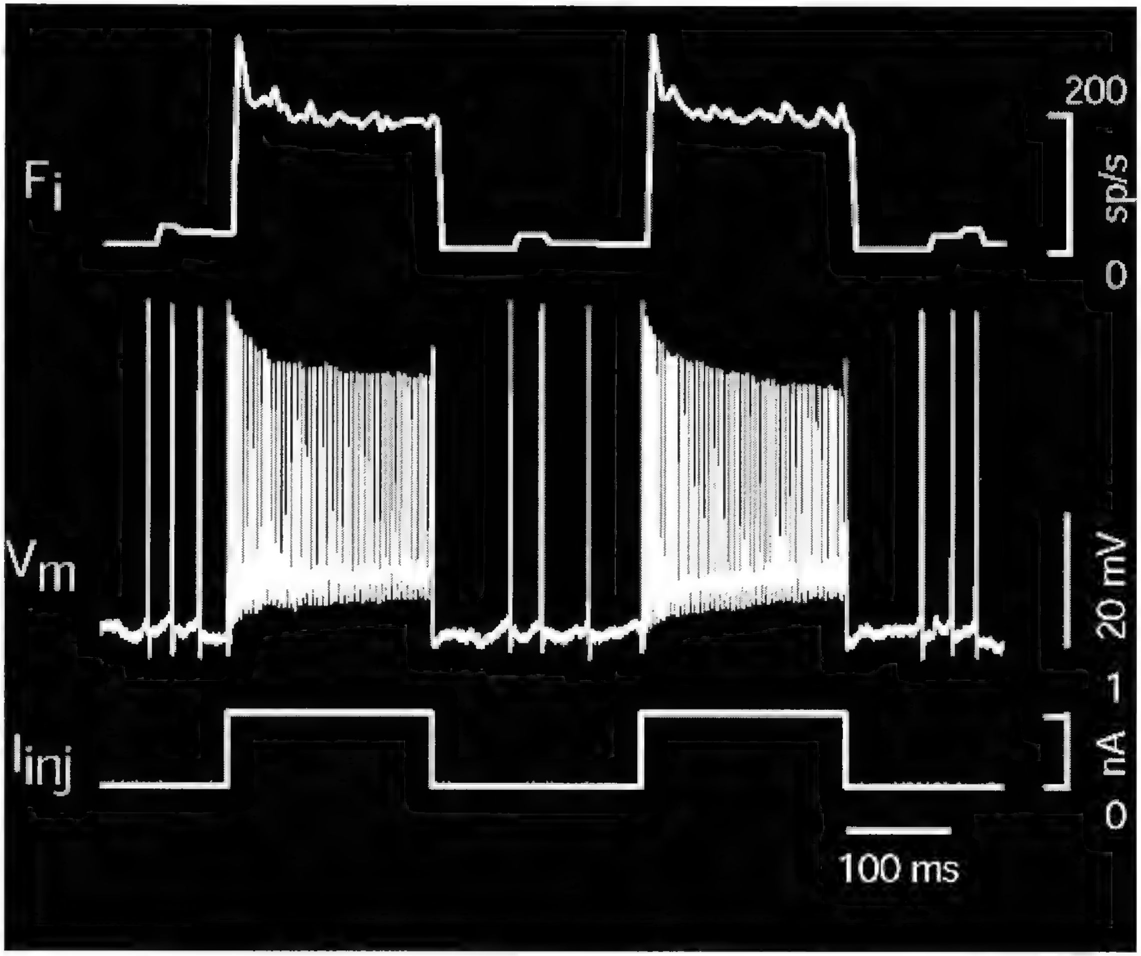
J. Hounsgaard and O. Kiehn
J. of Physiology 414, 265-282 (1989)

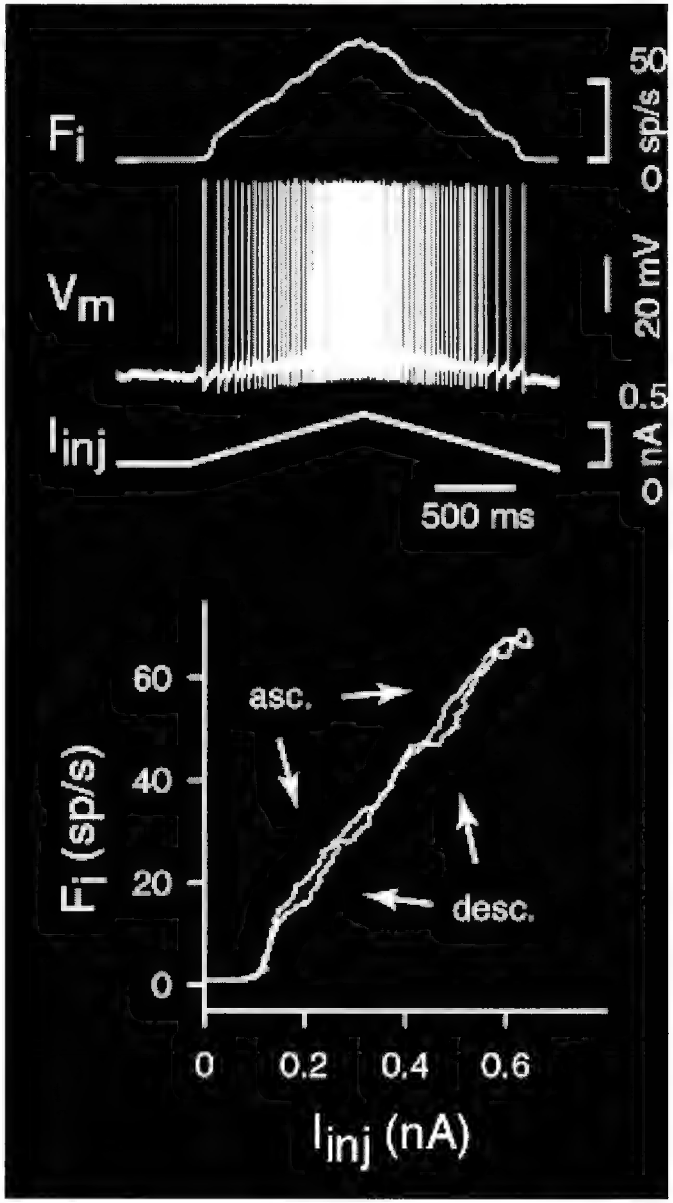
[redacted]

tiate

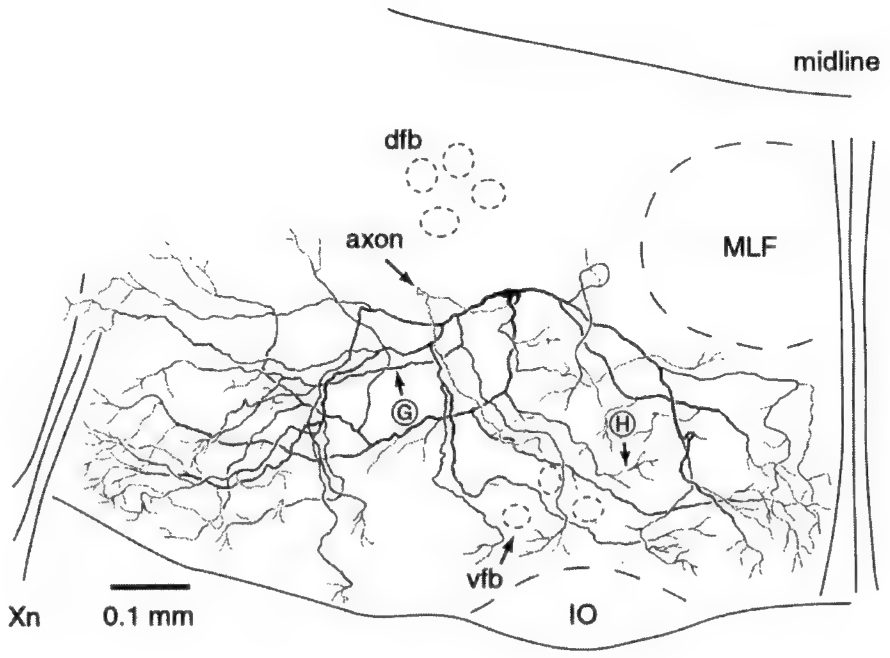
ulses

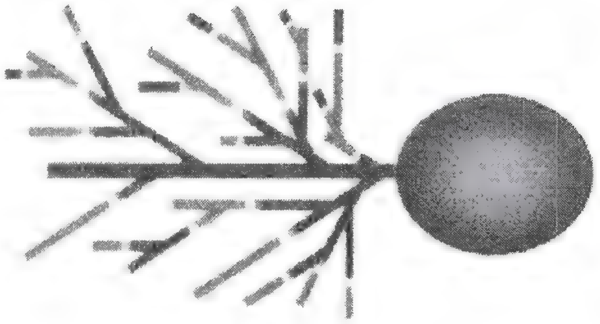
Test for Plateau Potentials



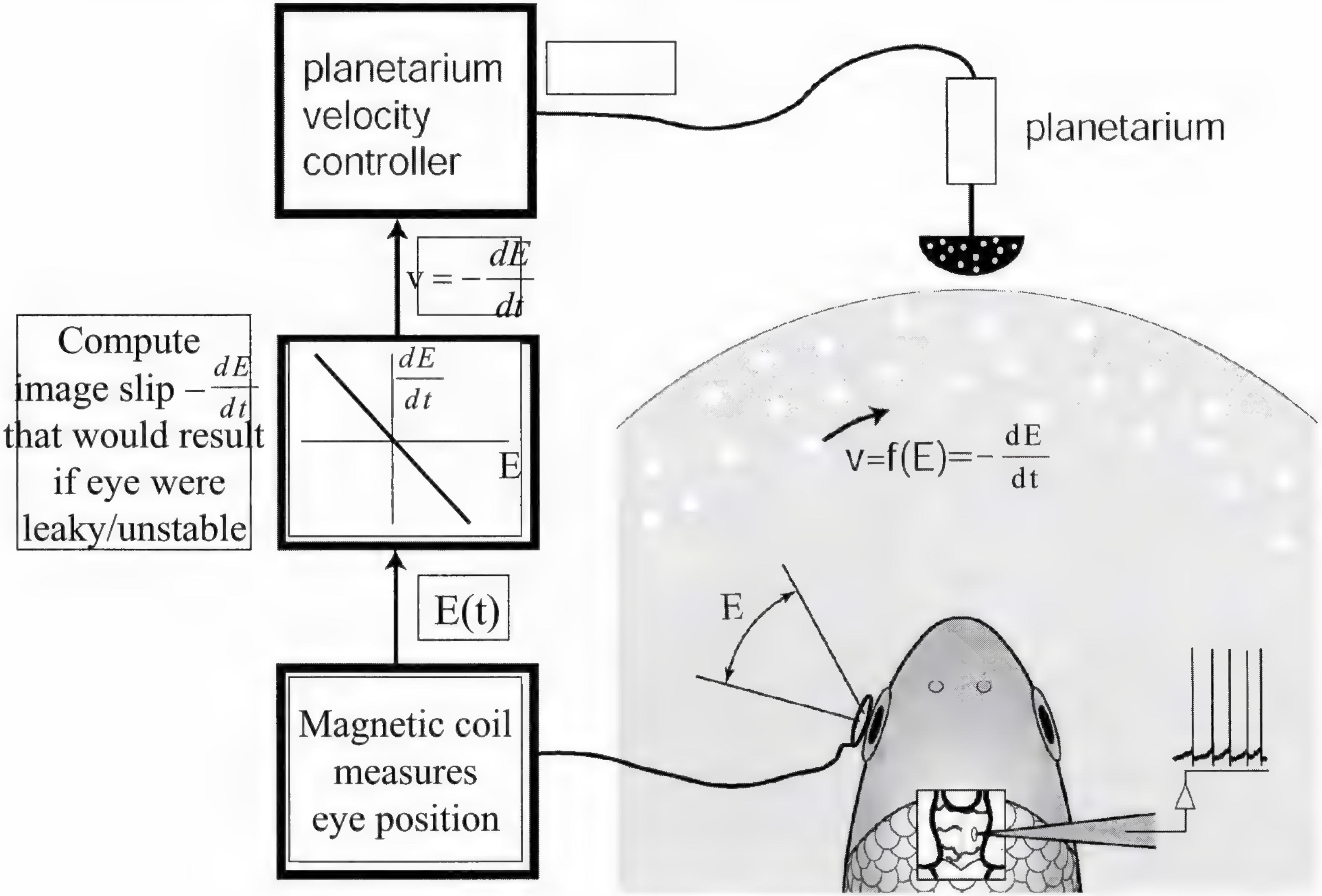


What is hiding in those dendrites?


midline
dfb
axon
MLF
G
H
vfb
Xn
0.1 mm
IO



Experiment: Give Feedback as if Eye is Leaky or Unstable



planetarium velocity controller

$v = -\frac{dE}{dt}$

Compute image slip $-\frac{dE}{dt}$ that would result if eye were leaky/unstable

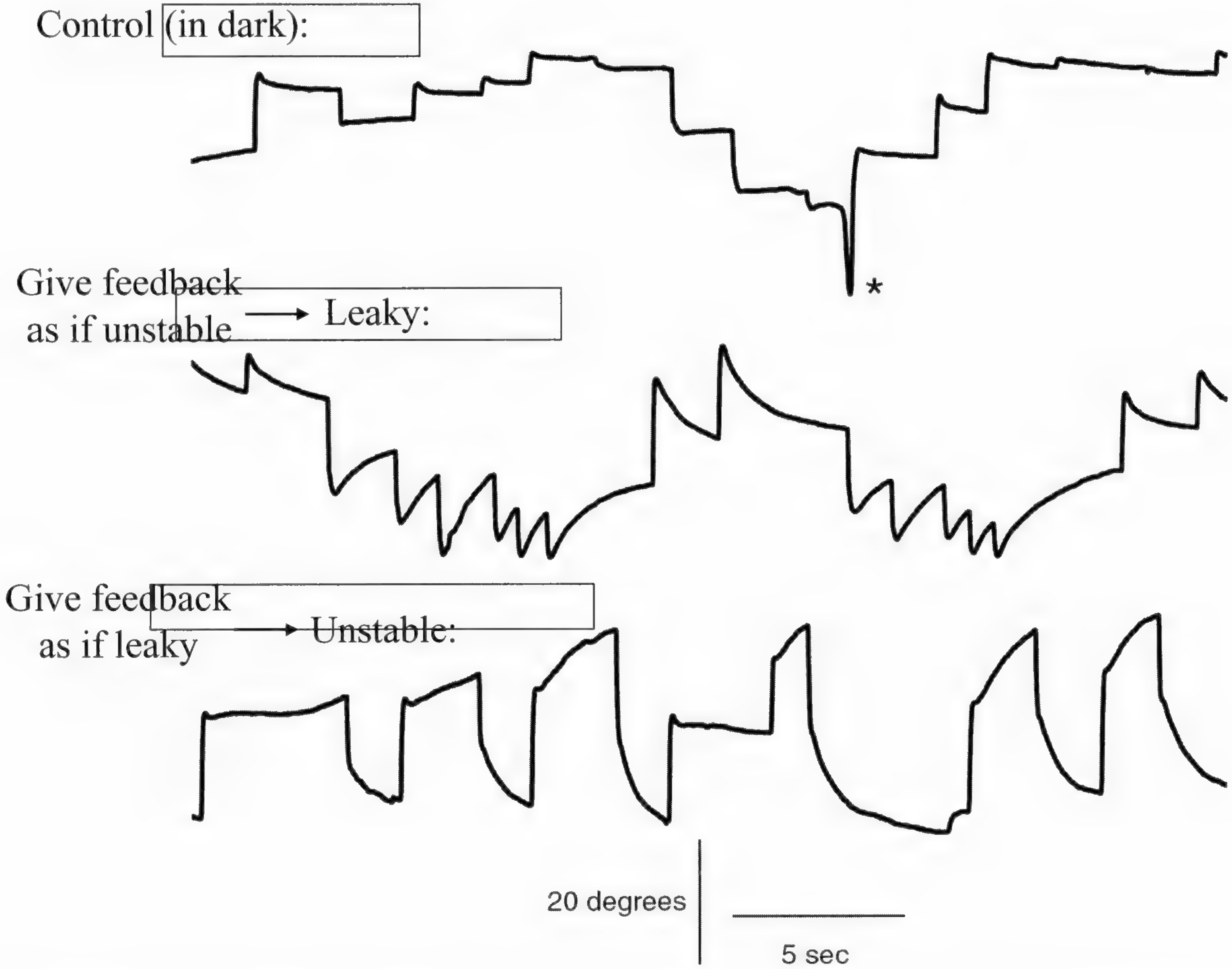
E(t)

Magnetic coil measures eye position

planetarium

$v = f(E) = -\frac{dE}{dt}$

Integrator Learns to Compensate for Leak/Instability!



Control (in dark):

Give feedback as if unstable → Leaky:

Give feedback as if leaky → Unstable:

Summary

Persistent neural activity: Activity which outlasts a transient stimulus

Neural integrator: Outputs the mathematical integral of its inputs
    -In the absence of input, maintains persistent neural activity

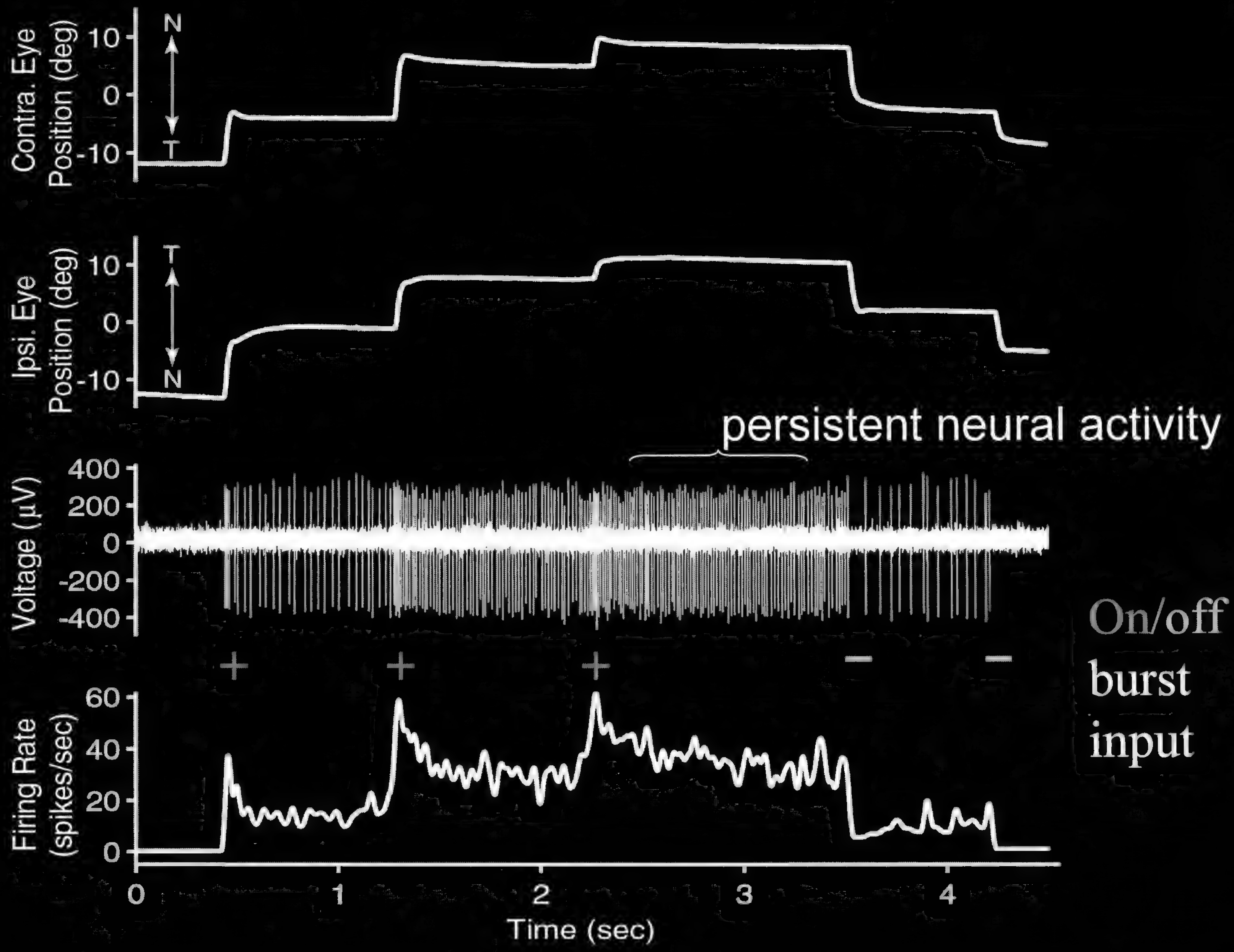
Oculomotor neural integrator:
    -Converts eye-velocity encoding inputs to eye position outputs
    -Mechanism:
        -Anatomy and physiology suggest network contribution
    -Issue of Robustness:
        Simple models are less robust to perturbations than the real system…
…But:
    -On fast time scales, intrinsic cellular processes may provide a “friction-like” slowing of network decay
    -Plasticity may keep the system tuned on slow time scales

Mechanisms and Models of Persistent Neural Activity: Linear Network Theory

Outline

1. Neural mechanisms of integration: Linear network theory
2. Critique of current models of memory-related activity, and possible remedies

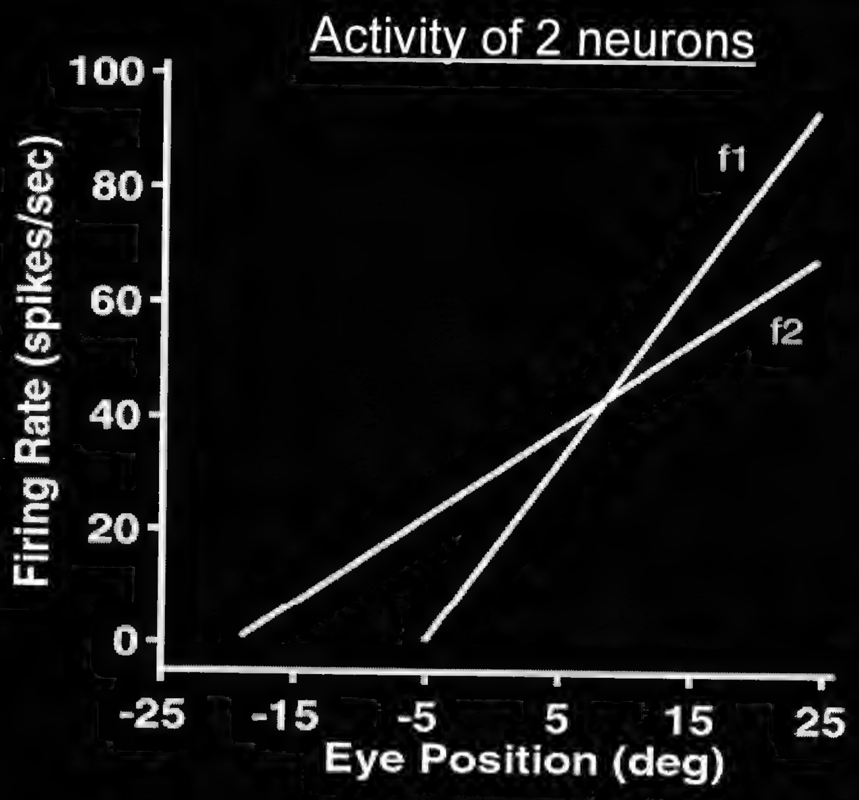
Neural Recording from the Oculomotor Integrator

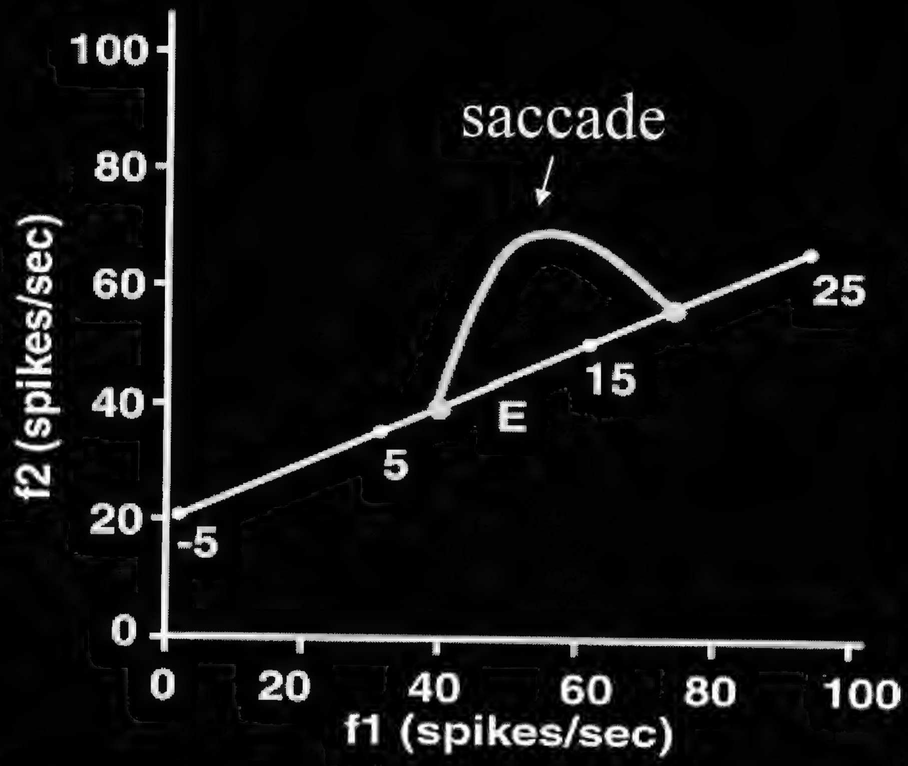


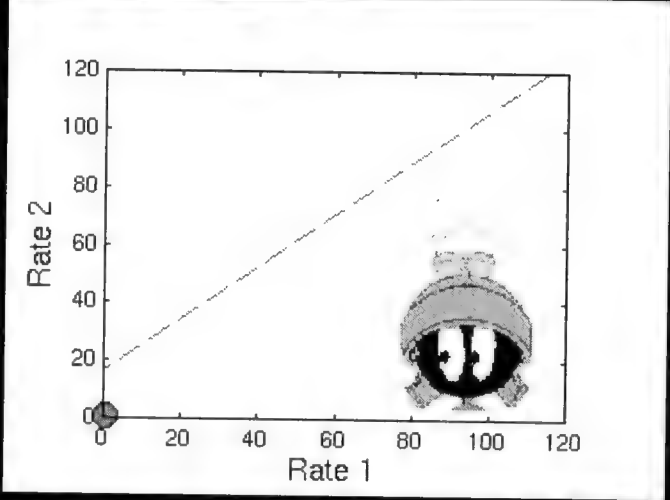
persistent neural activity

On/off
burst
input

Many-neuron Patterns of Activity Represent Eye Position


Activity of 2 neurons



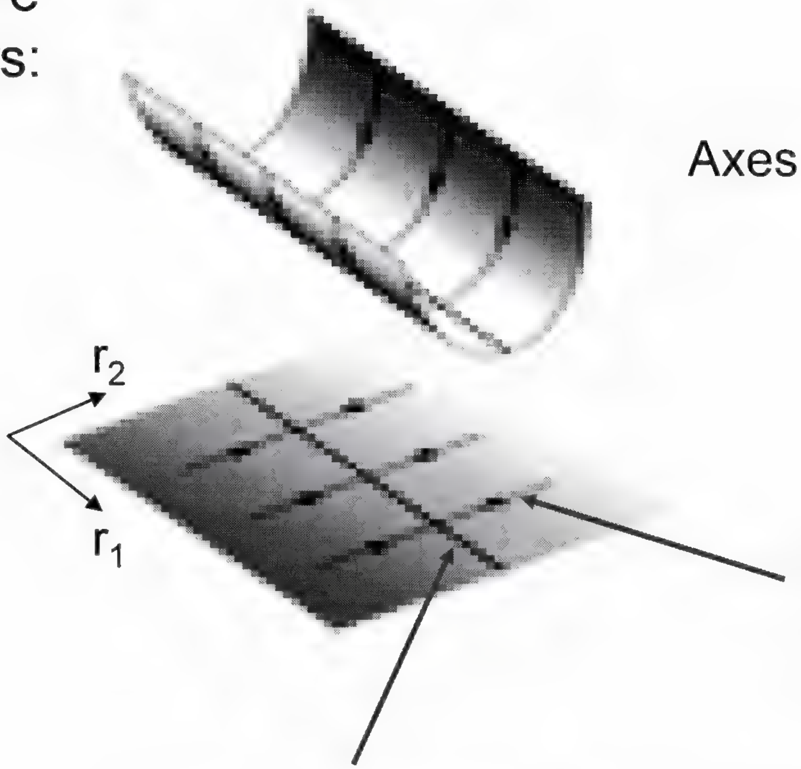


Eye position represented
by location along
a low dimensional
manifold ("line attractor")

(H.S. Seung, D. Lee)

Line Attractor Picture of the Neural Integrator

Geometrical picture
of eigenvectors:



Axes = firing rates
of 2 neurons

Decay along direction
of decaying eigenvectors

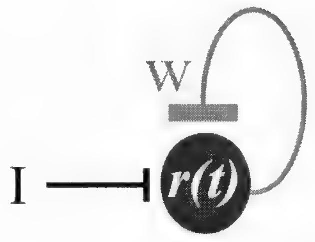
No decay or growth along direction
of eigenvector with eigenvalue = 1

"Line Attractor" or "Line of Fixed Points"

Issue: Robustness of Integrator

Integrator equation:

$$\tau_{bio} \frac{dr}{dt} = -1r + wr + I$$



$$\tau_{network} = \frac{\tau_{bio}}{|1 - \mathrm{w}|}$$

Experimental values:

Single isolated neuron: $\tau_{bio} \sim 100\ \mathrm{ms}$

Integrator circuit: $\tau_{network} \sim 30\ \mathrm{sec}$

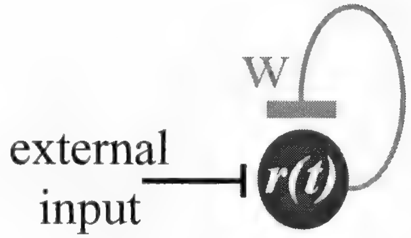
➡ Synaptic feedback w must be tuned to accuracy of:

$$|1 - \mathrm{w}| = \frac{\tau_{bio}}{\tau_{network}} \sim 0.3\%$$

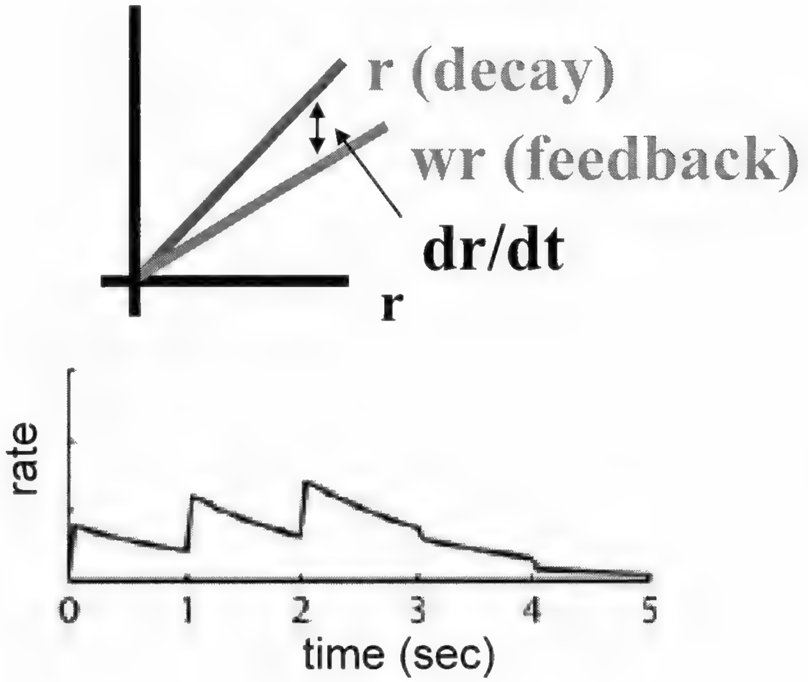
Need for fine-tuning in linear feedback models

Fine-tuned model:

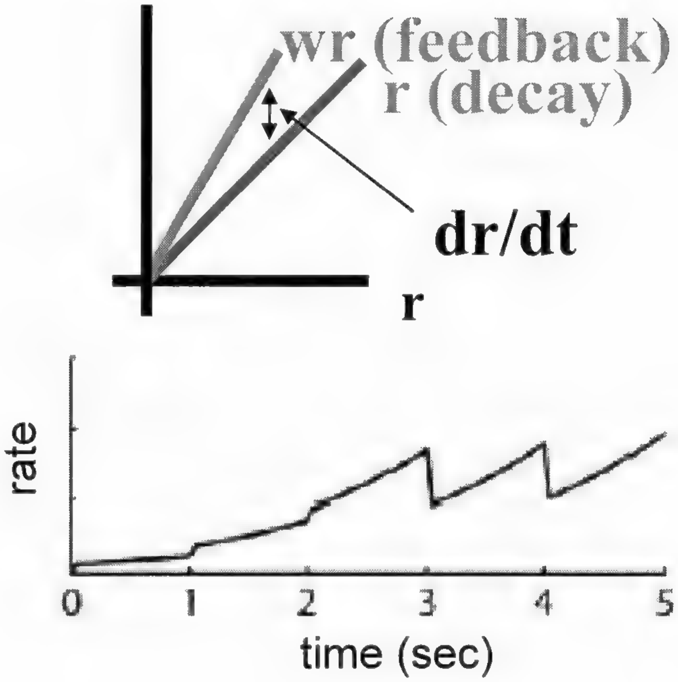
$$\tau_{neuron}\frac{dr}{dt} = \underbrace{-r}_{\text{decay}} + \underbrace{\mathrm{w}r}_{\text{feedback}} + external\ input$$



Leaky behavior



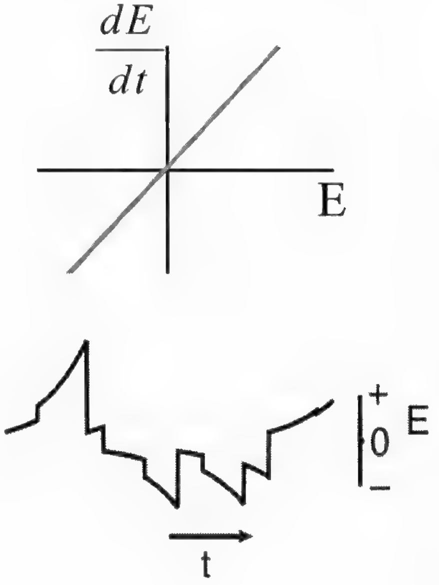
Unstable behavior



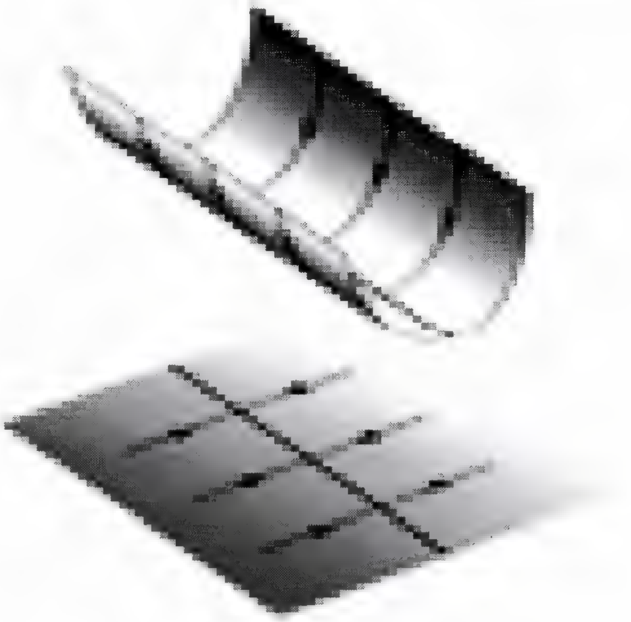
Unstable Integrator

Human with unstable integrator:

$$\frac{dE}{dt} = (w-1)E$$

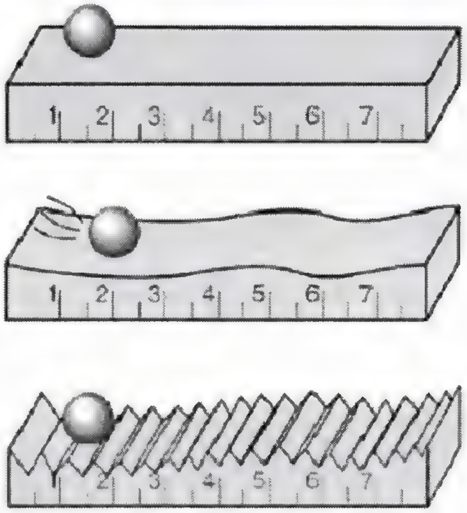

$\frac{dE}{dt}$
E
+
0
–
E
t

Geometry of Robustness
& Hypothesis for Robustness on Faster Time Scales



1) Plasticity on slow time scales:
Reshapes the trough to make it flat

2) To control on faster time scales:
Add ridges to surface to add
“friction”-like slowing of drift

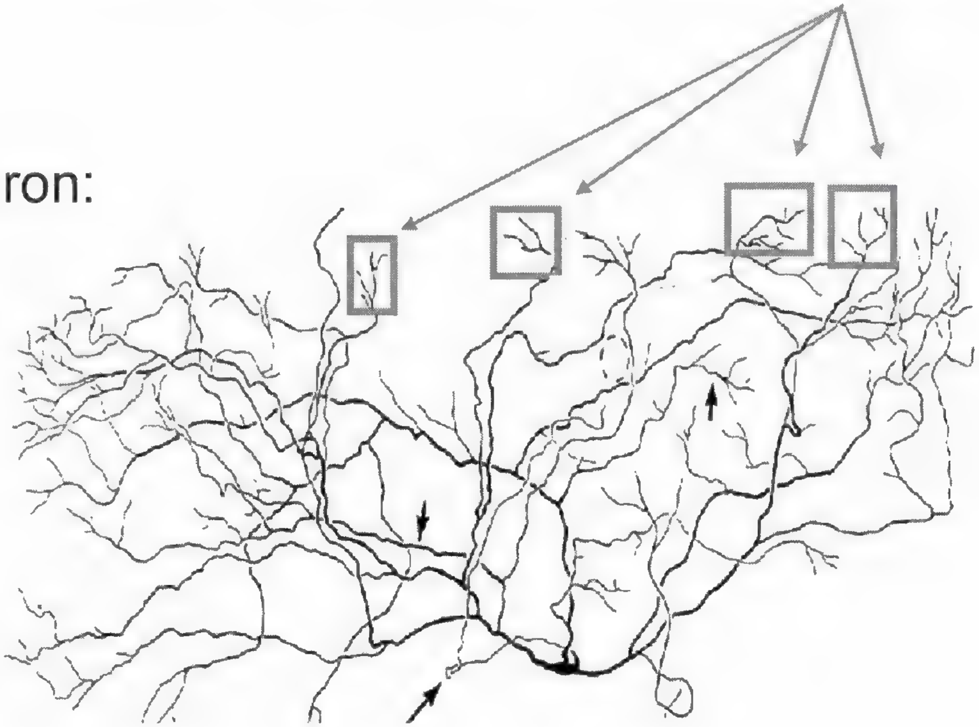


1 possible solution to robustness problem: Bistable dendrites

Idea: Neurons may have intrinsic properties that help to maintain the persistent neural activity

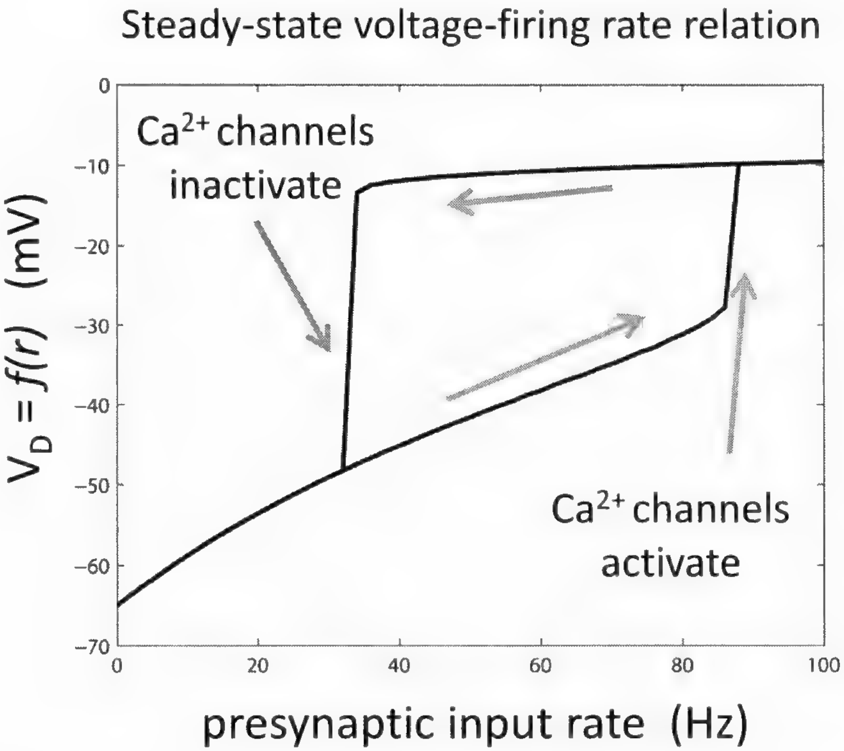
- ❖ Possible intrinsic mechanism:
  Dendritic bistability distributed across multiple independent dendrites

Integrator neuron:



Evidence for dendritic bistability & independence

1) Dendritic bistability has been observed experimentally
   - Due to the *self-sustaining* properties of NMDA, NaP, or Ca^{++} channels:

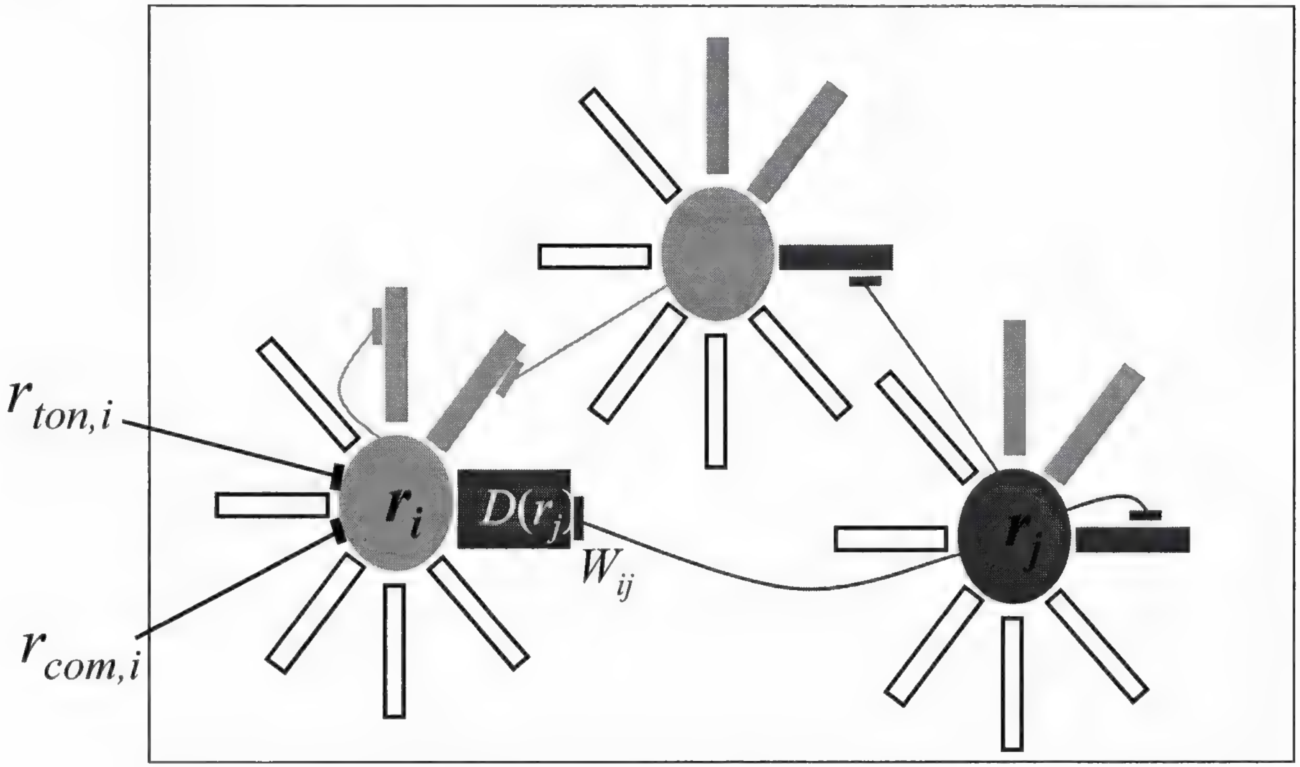


(Lee, Aksay, & Goldman, in preparation)

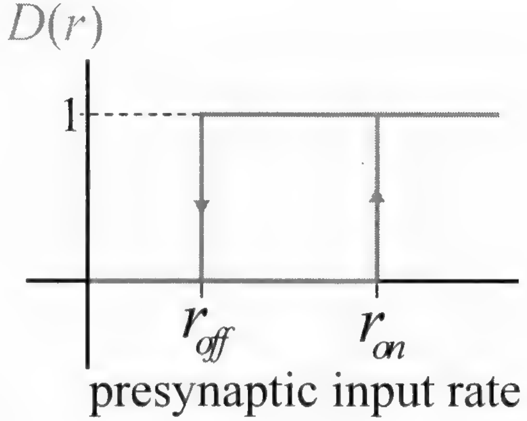
2) Anatomically realistic models suggest that different dendritic branches may behave approximately independently (Koch et al., 1983; Poirazi et al., 2003)

Network with Bistable Dendrites

Network of *N* neurons, each with *N* identical dendrites:



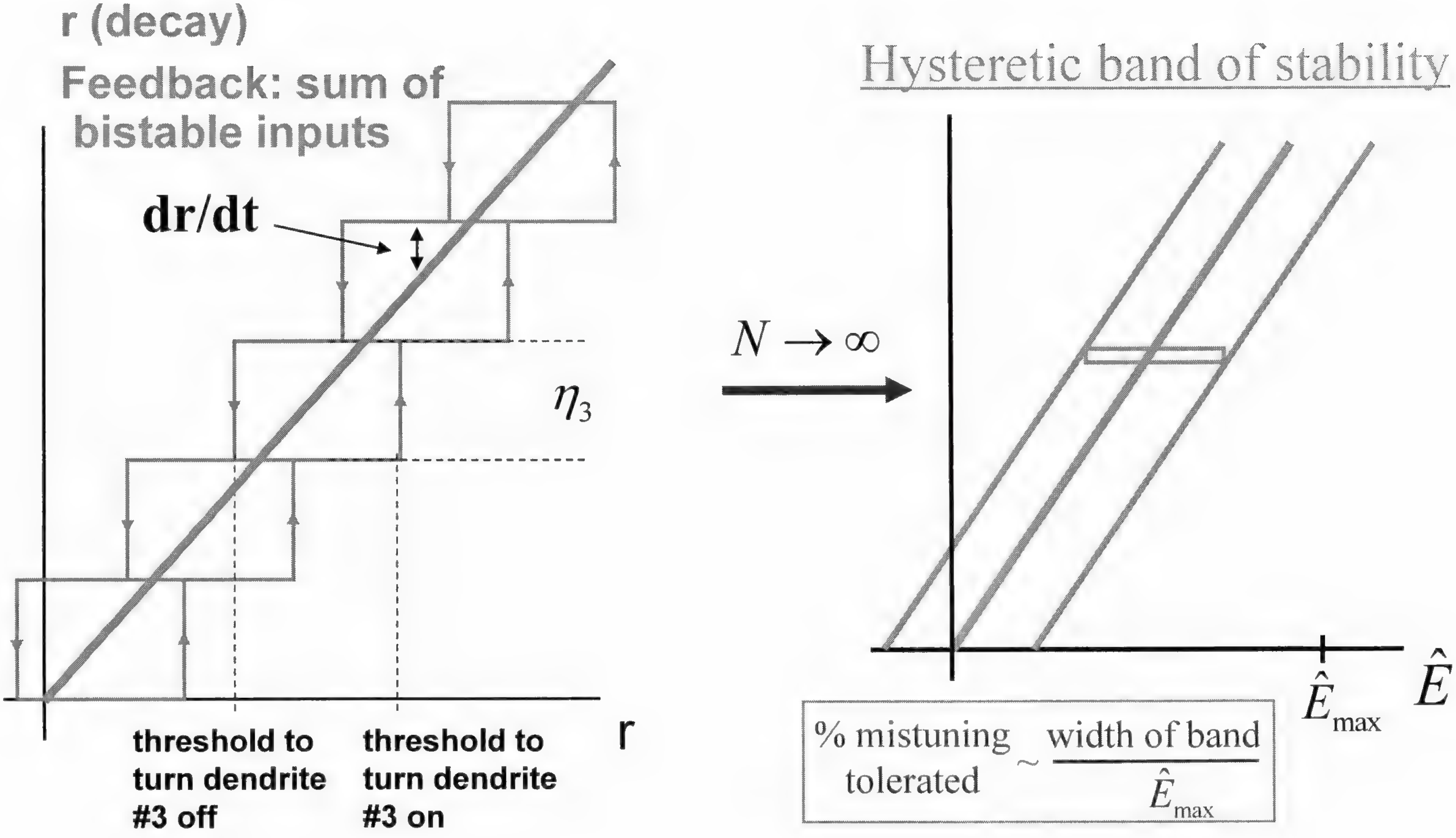
Bistable dendritic response



$$\tau \frac{dr_i}{dt} = \underbrace{-r_i}_{\text{decay}} + \underbrace{\sum_{j=1}^{N} W_{ij} D(r_j)}_{\text{recurrent input}} + \underbrace{r_{ton,i}}_{\text{background input}} + \underbrace{r_{com,i}}_{\text{eye movement commands}}$$

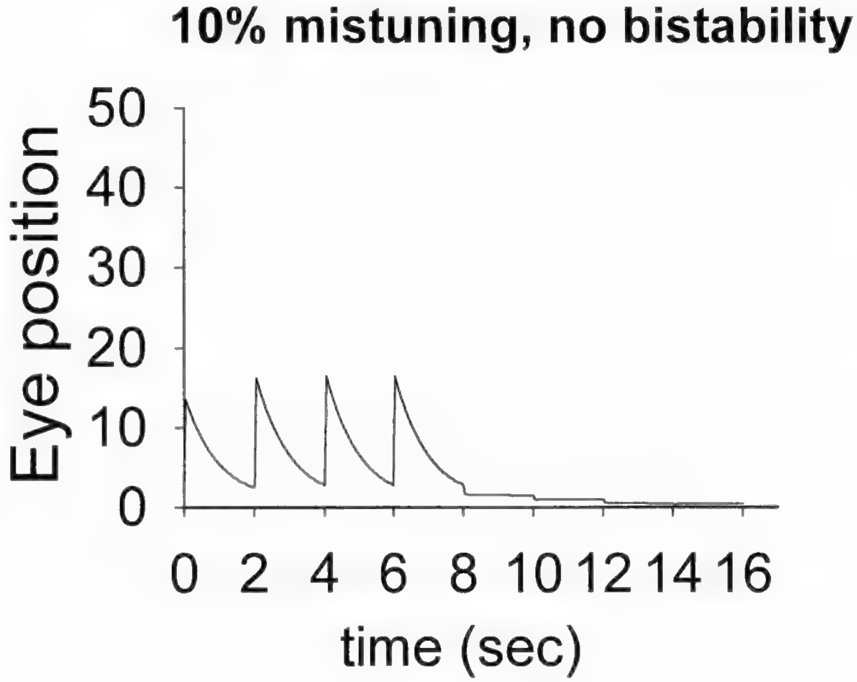
Graphical solution of balanced leak and feedback

No external input:

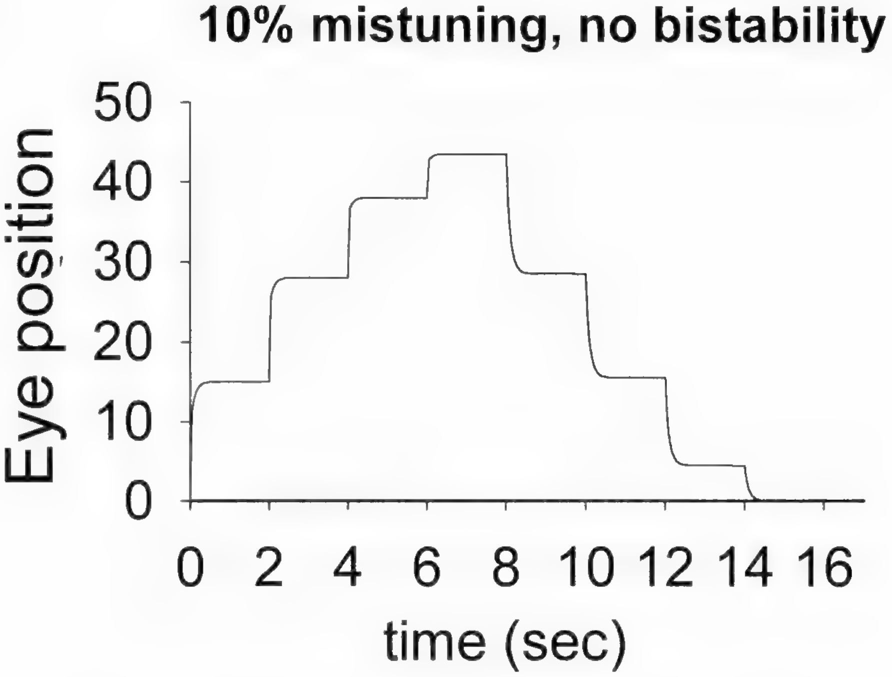

r (decay)
Feedback: sum of bistable inputs
dr/dt
η_3
threshold to turn dendrite #3 off
threshold to turn dendrite #3 on
r
$N \to \infty$
Hysteretic band of stability
$\hat{E}_{\max}$
$\hat{E}$

% mistuning tolerated $\sim \frac{\text{width of band}}{\hat{E}_{\max}}$

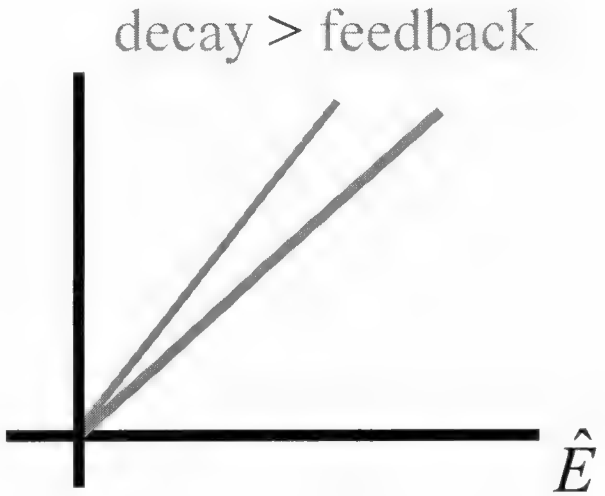
Comparison of robustness with & without bistability



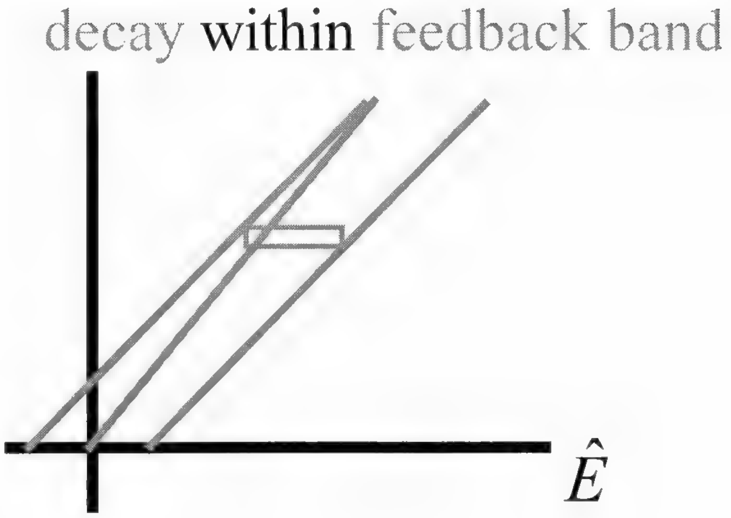
10% mistuning, no bistability



10% mistuning, no bistability

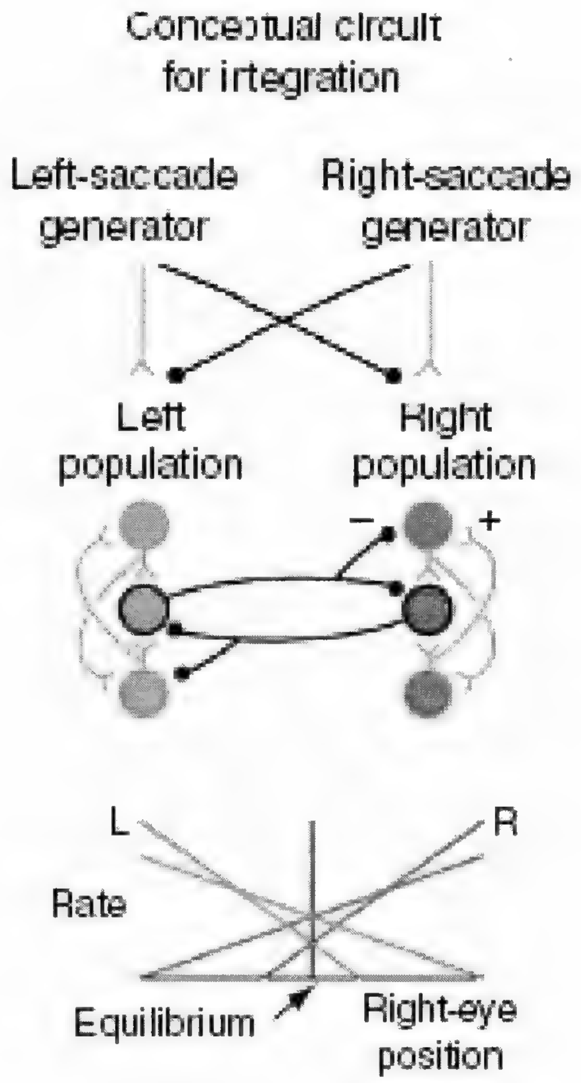


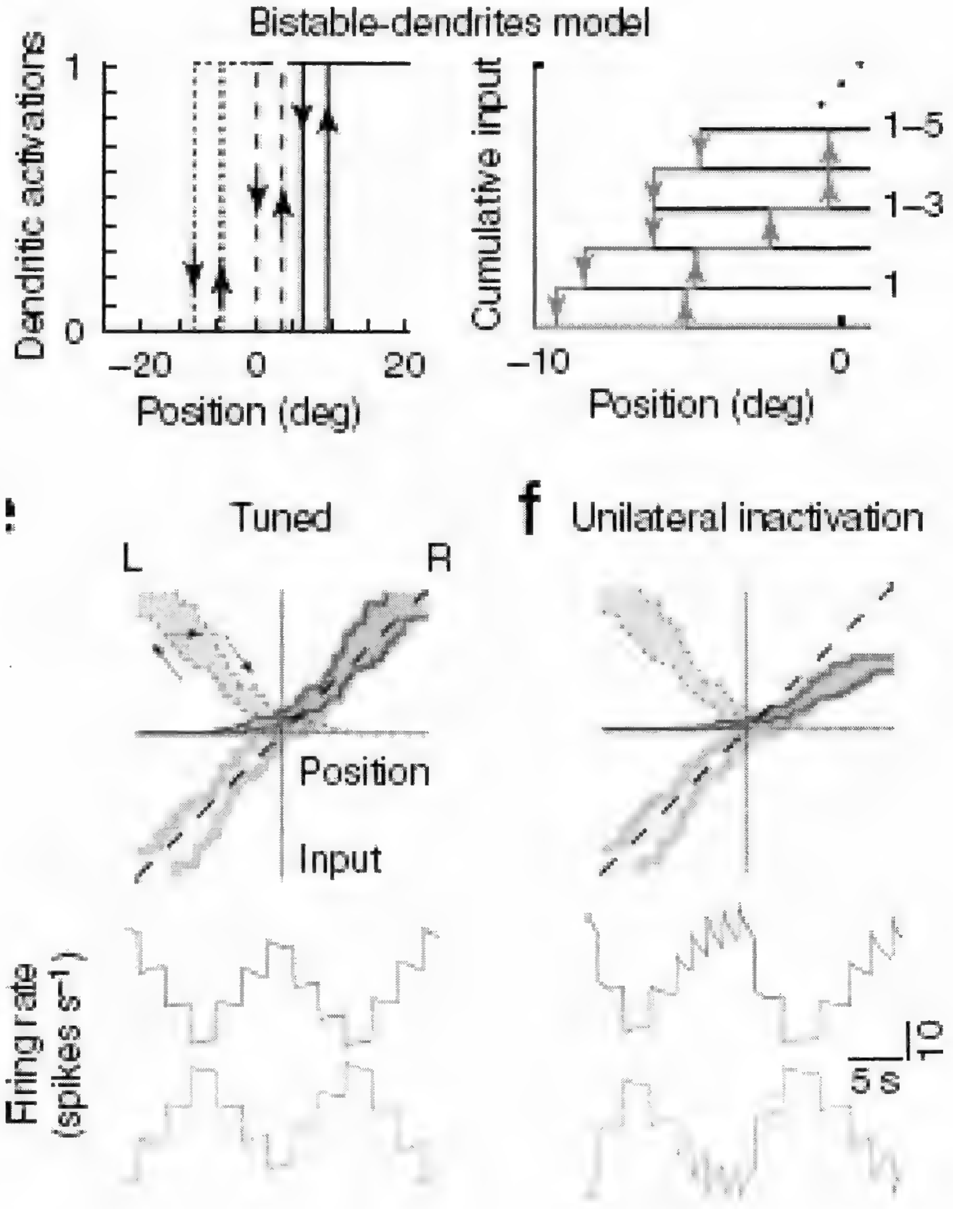
decay > feedback
$\hat{E}$



decay within feedback band
$\hat{E}$

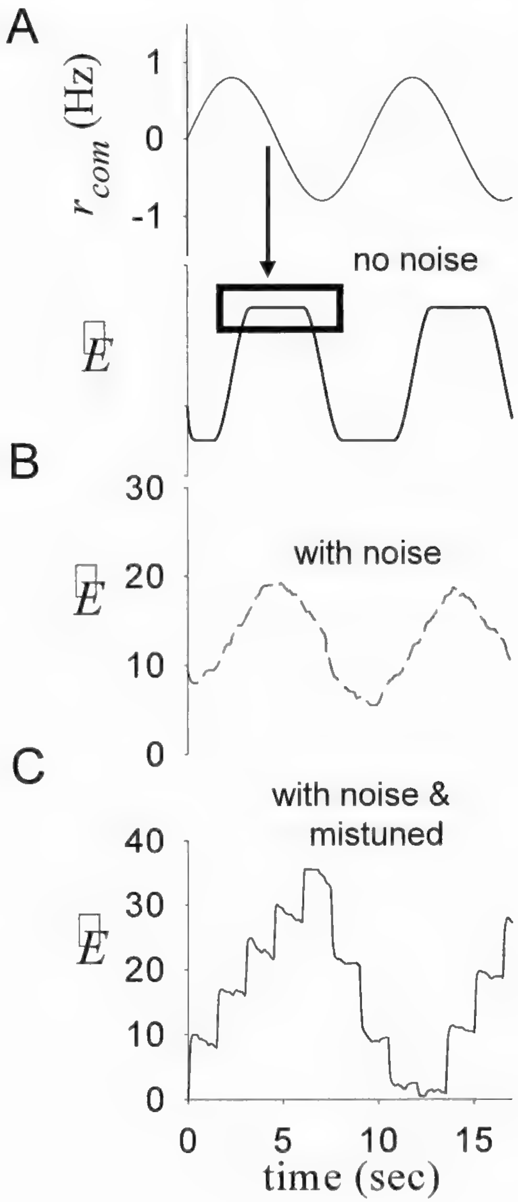
Bilateral Model with Realistic Neuronal Firing Rates


Conceptual circuit
for integration
Left-saccade generator
Right-saccade generator
Left population
Right population
−
+
L
R
Rate
Equilibrium
Right-eye position



(Aksay et al., Nature Neuroscience, 2007)

Key issues with robustness through bistability: Insensitivity to small inputs

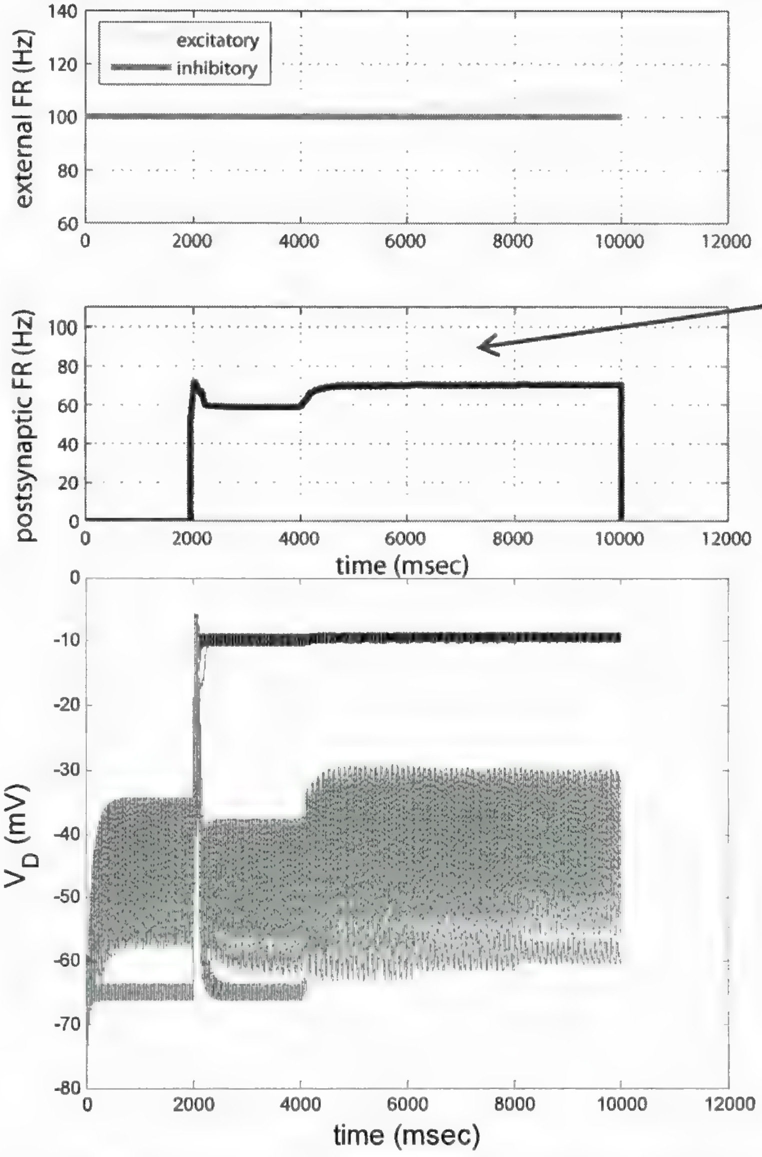


Goldman et al, Cerebral Cortex, 2003

Biophysical Model

(Lee, Aksay, & Goldman, in preparation)

Insensitivity to weak stimuli

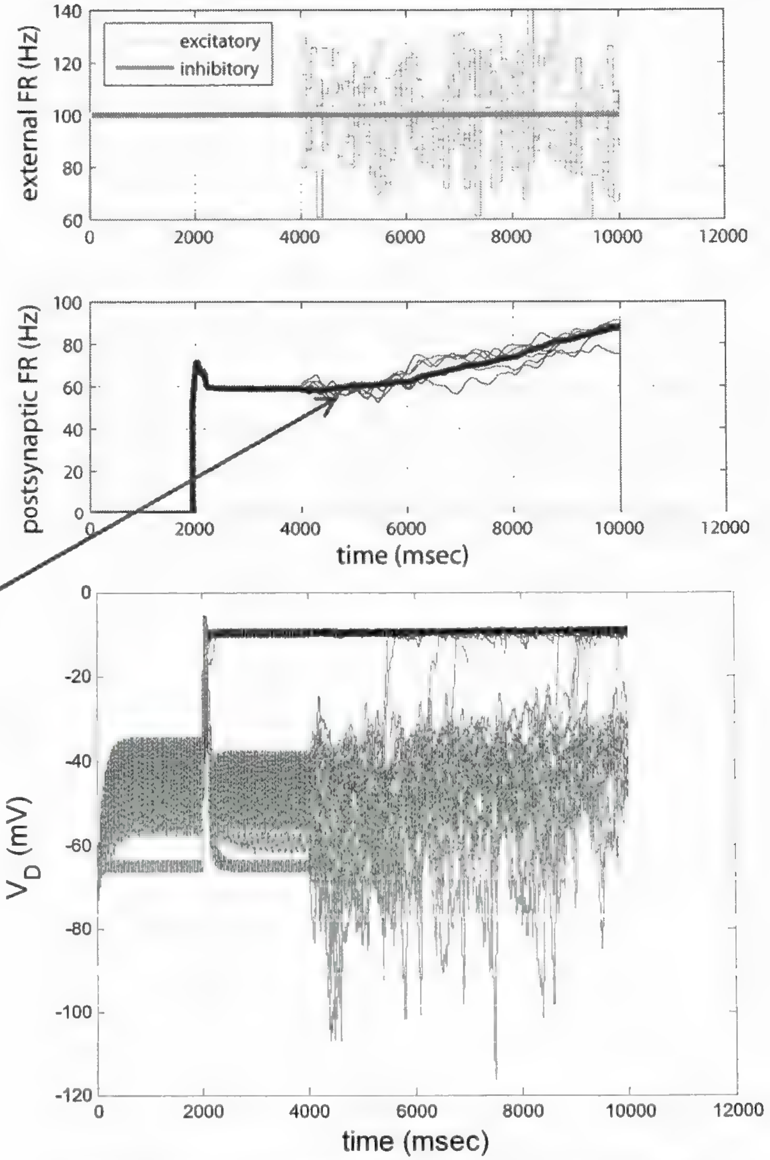


Dendrites do not change their activation state due to bistability/insensitivity to small inputs

2 Hz steps in input FR are not integrated in FR response

2 Hz steps in FR are integrated to yield ramps of increasing slope (bold line: average of 100 runs)

Sensitivity regained with noise



Dendrites change activation states with noisy input

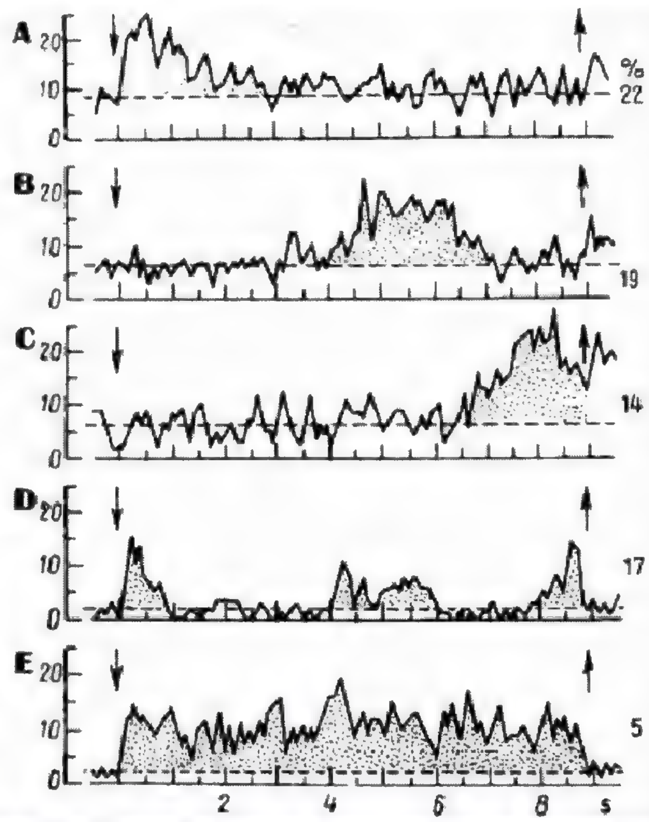
Questions:

1) Are positive feedback loops the only way to perform integration? (the dogma)

2) Could alternative mechanisms describe persistent activity data?

Working memory task not easily explained by traditional feedback models

5 neurons recorded during a PFC delay task (Batuev et al., 1979, 1994):



Early stage neuron?

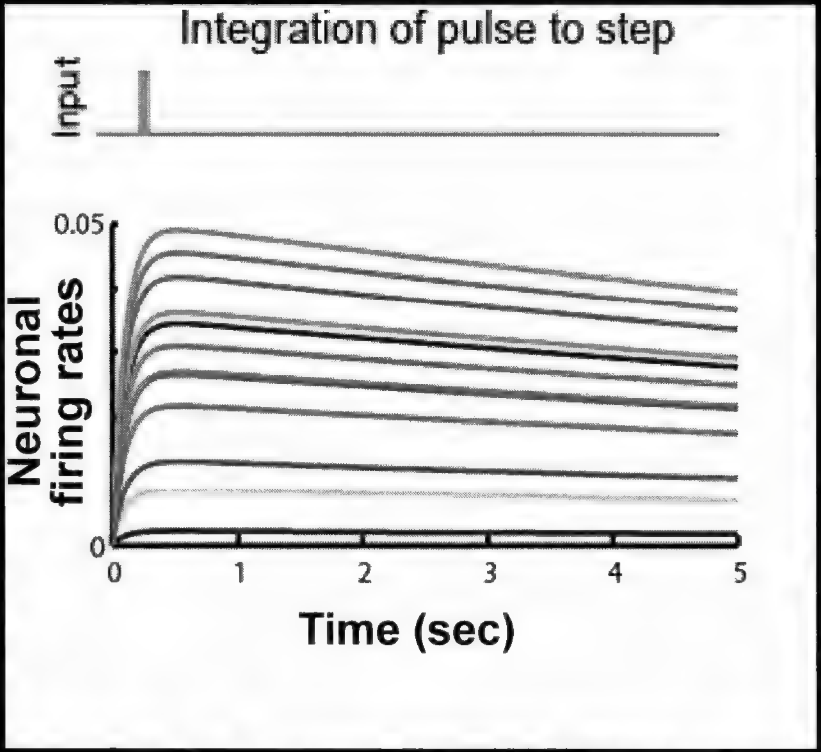
Middle stage neuron?

Late stage neuron?

Sum early, mid, late
stage neurons?

Sum all stages?

Response of Individual Neurons in Line Attractor Networks



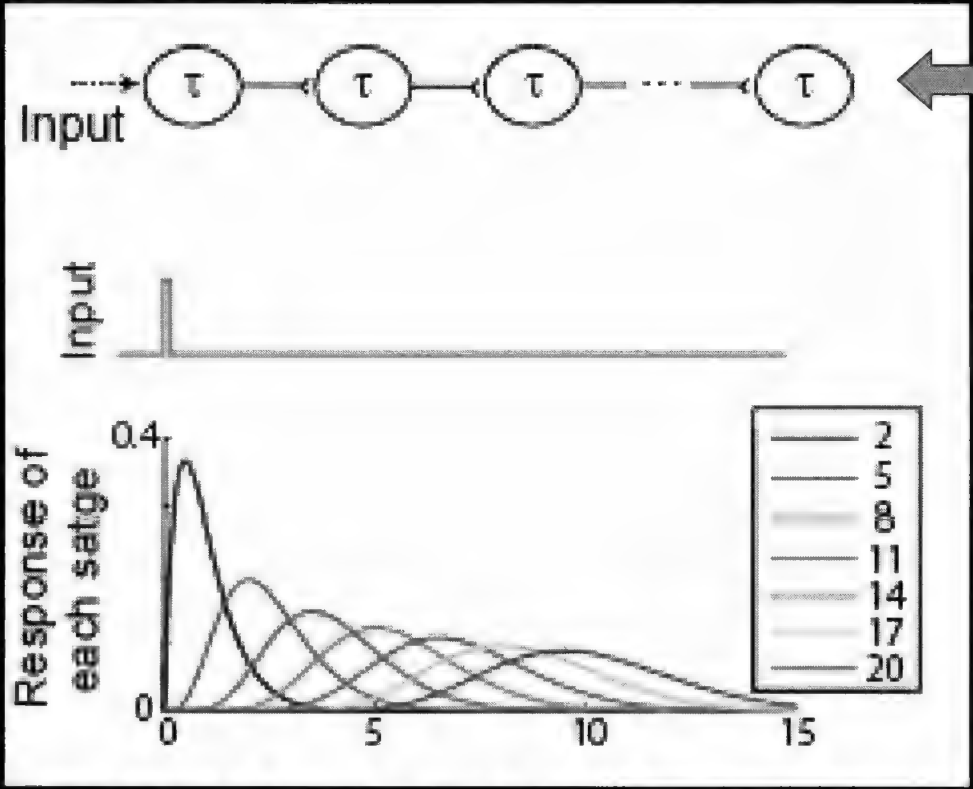
All neurons exhibit similar slow decay:
Due to strong coupling that mediates positive feedback

Problem: Does not reproduce the differences between neurons seen experimentally!

Feedforward Networks Can Integrate!

(Goldman, *Neuron*, 2009)

Simplest example:

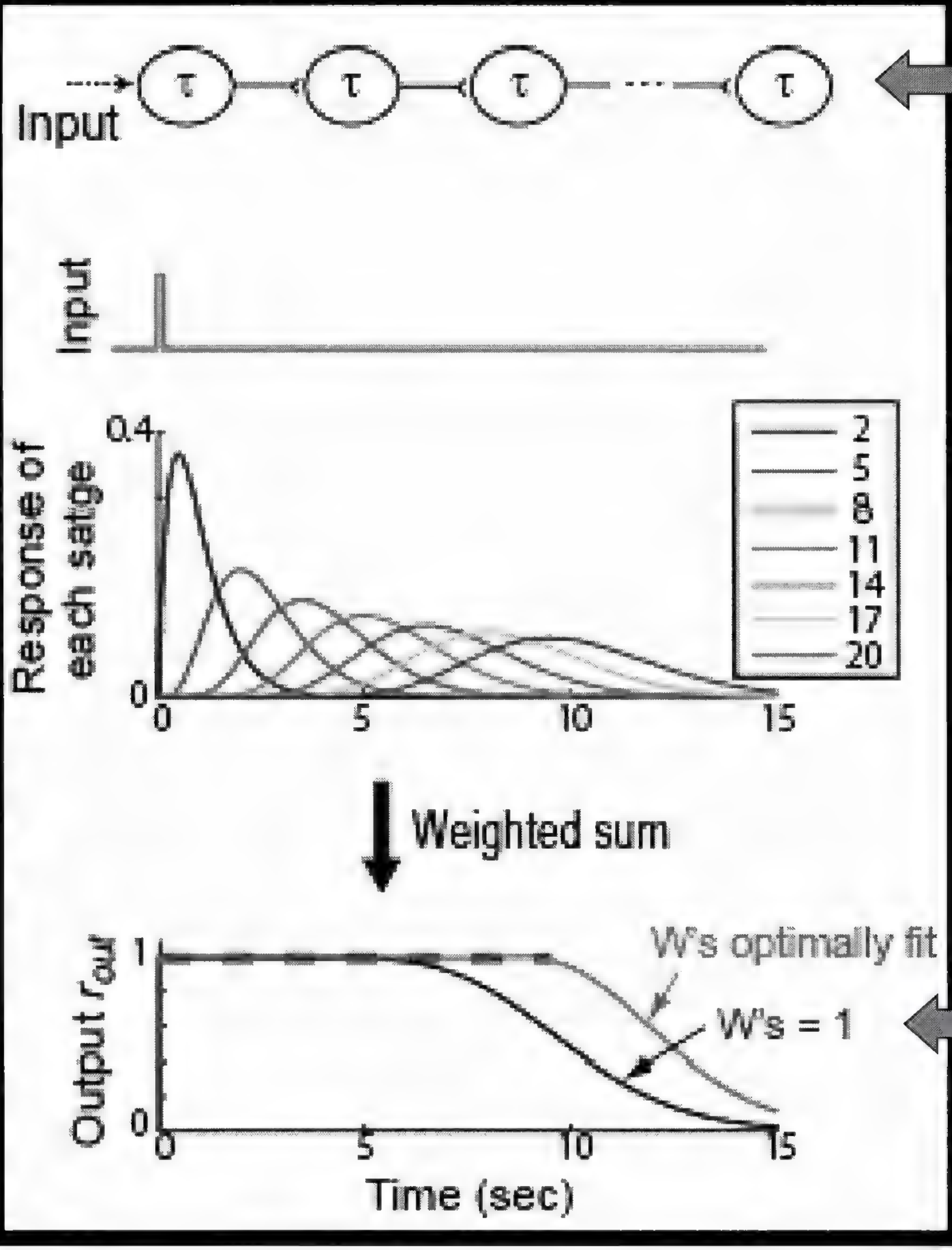


Chain of neuron clusters that successively filter an input

Feedforward Networks Can Integrate!

(Goldman, *Neuron*, 2009)

Simplest example:



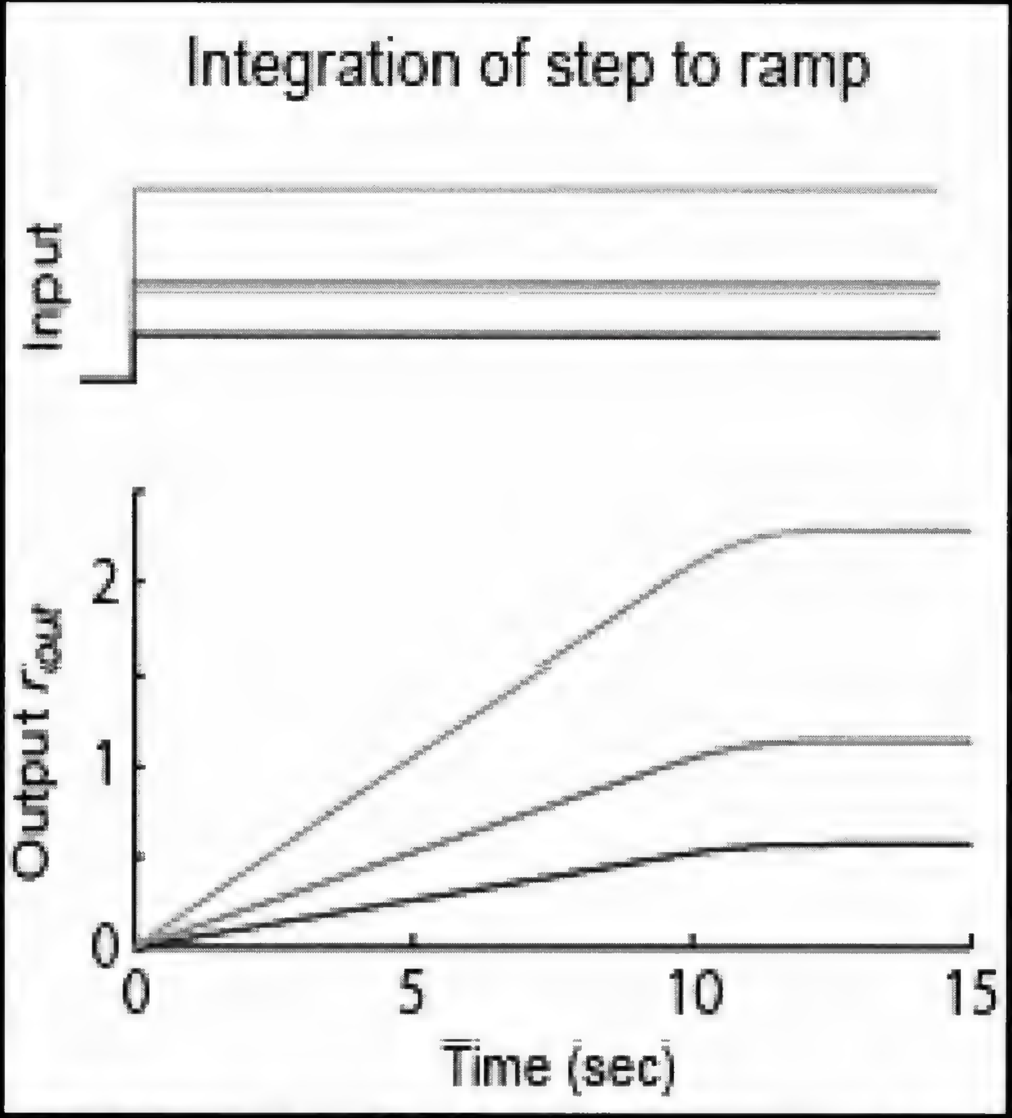
Chain of neuron clusters that successively filter an input

Integral of input!

(up to duration ~Nτ)

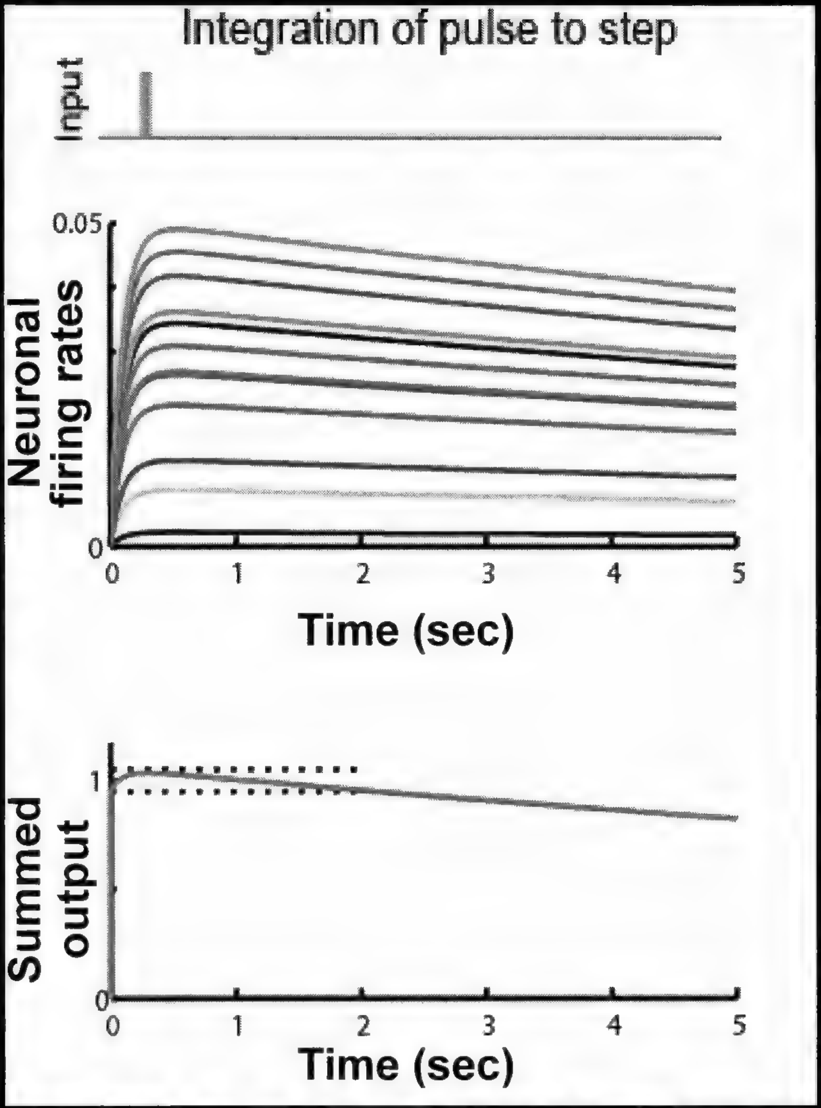
(can prove this works analytically)

Same Network Integrates Any Input for ~Nτ

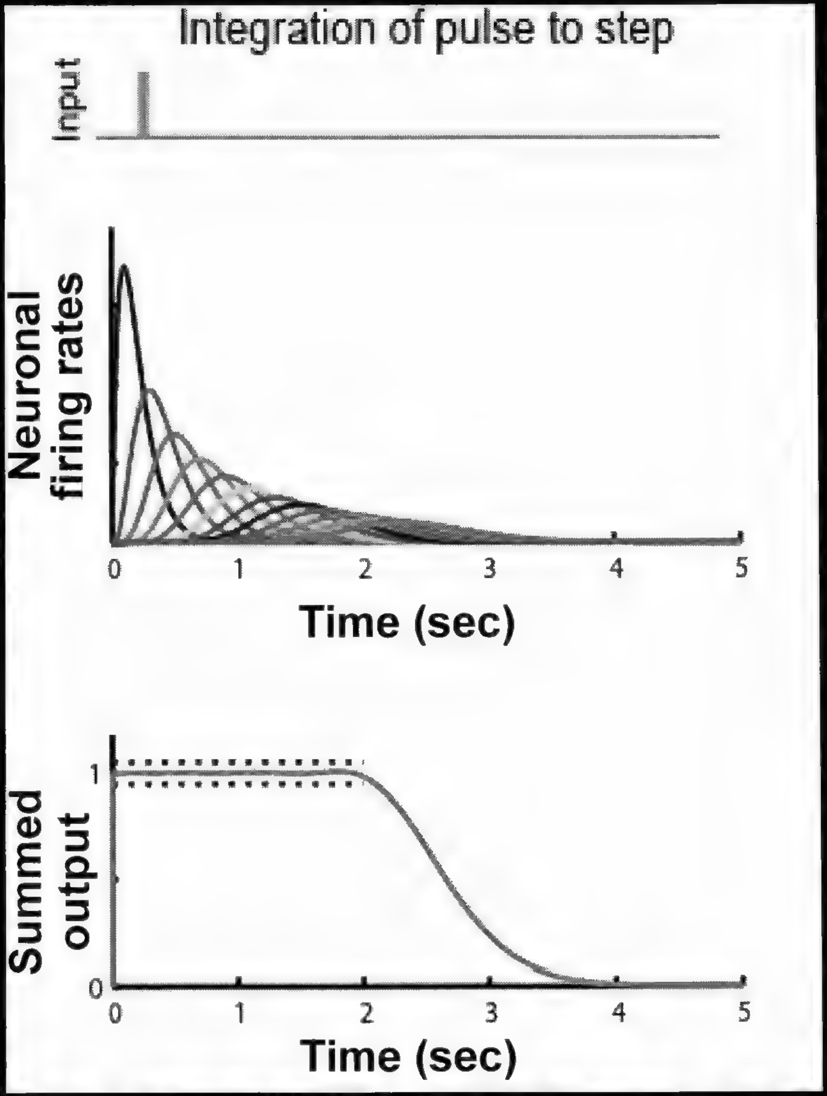


Improvement in Required Precision of Tuning

Feedback-based Line Attractor:
10 sec decay to hold 2 sec of activity

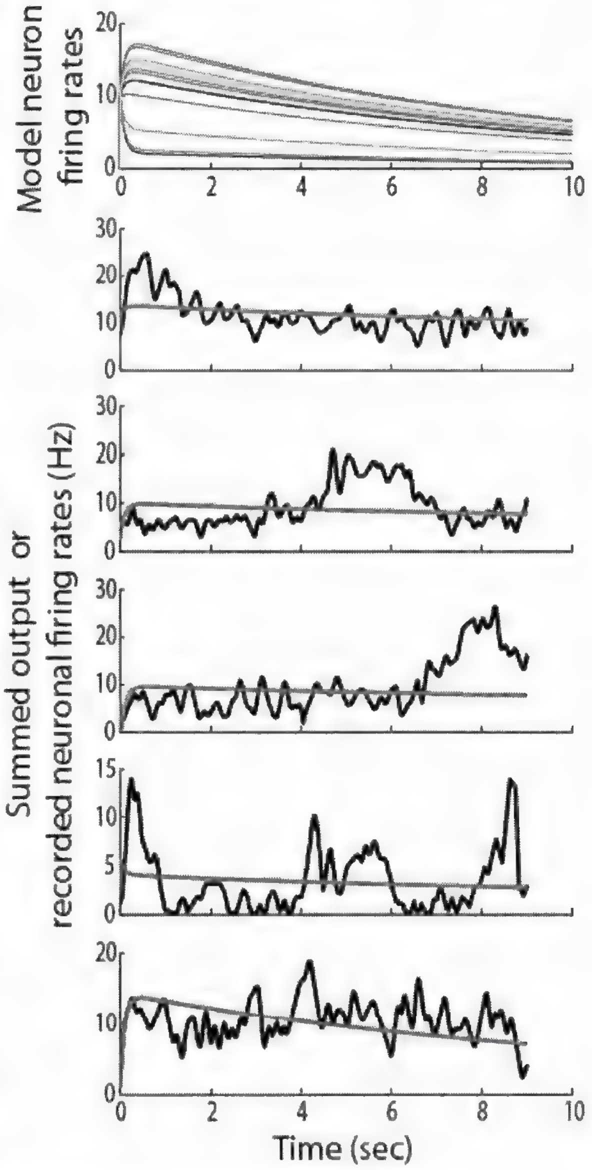


Feedforward Integrator
2 sec decay to hold 2 sec of activity

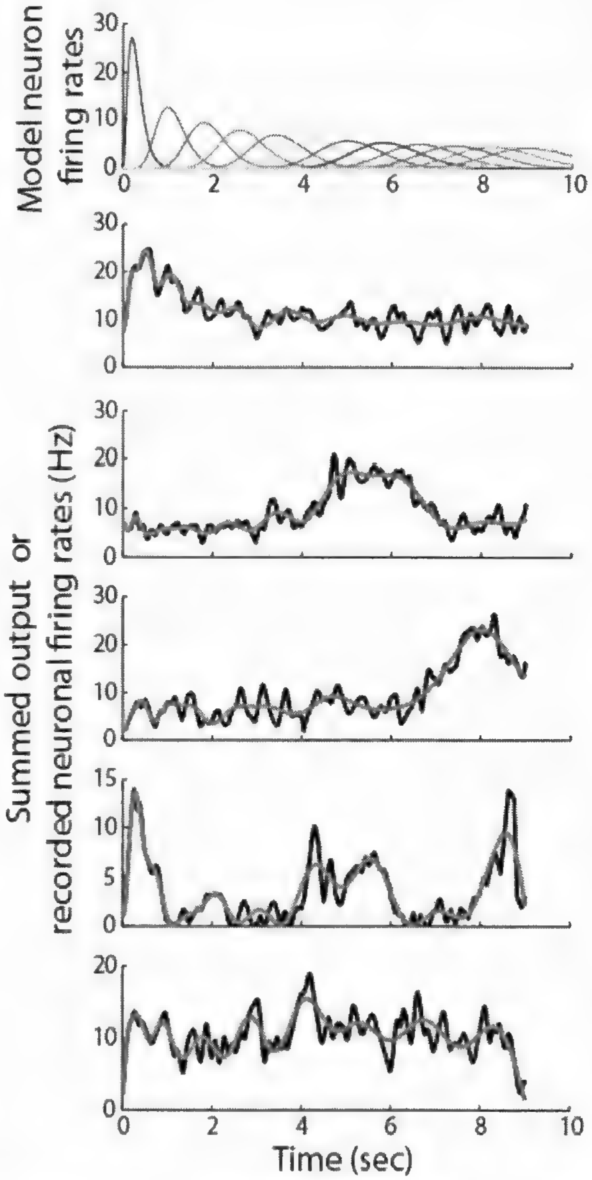


Feedforward Models Can Fit PFC Recordings

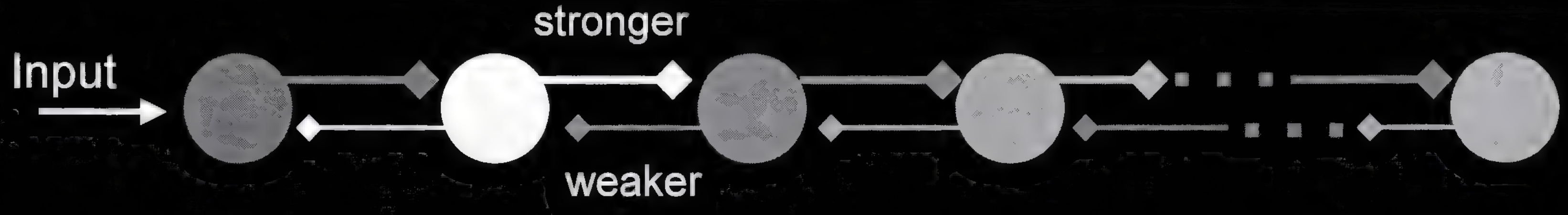
Line Attractor



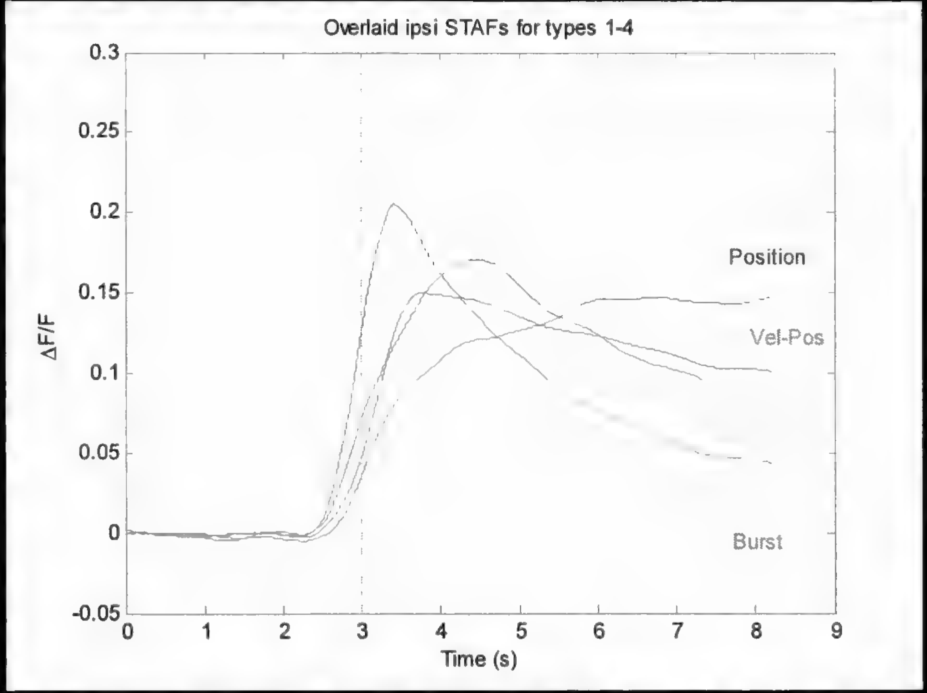
Feedforward Network



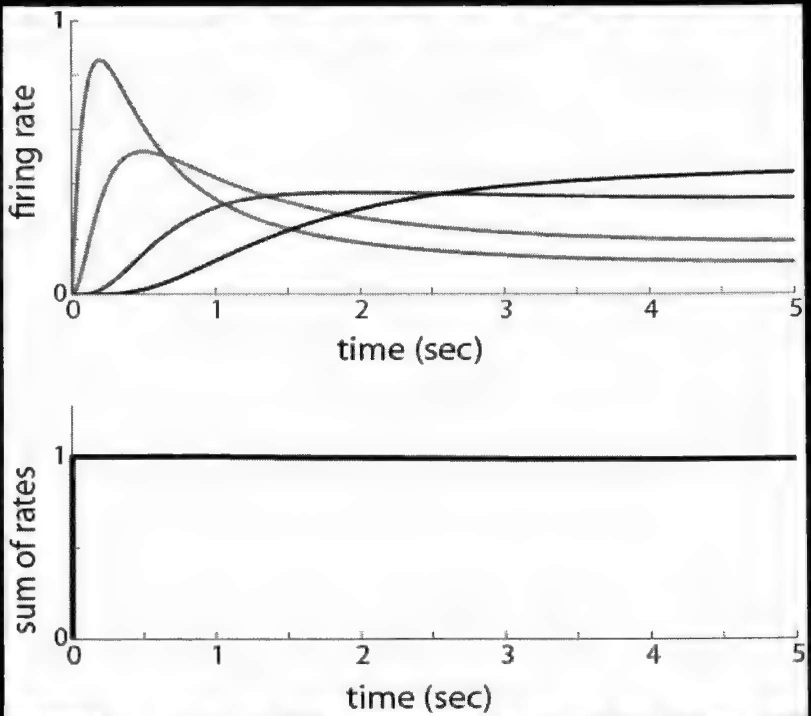
Crude asymmetric network model of Aksay et al. imaging data


Input
stronger
weaker

Larval zebrafish experiment

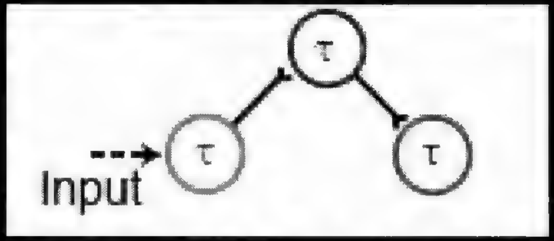


Crude model



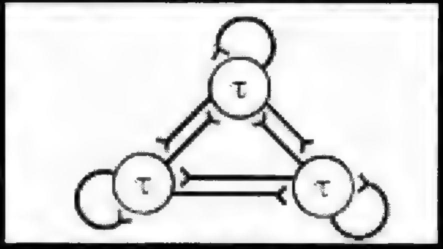
Generalization to Coupled Networks: Feedforward transitions between *patterns* of activity

Feedforward
network



Connectivity
matrix W_{ij}: $\mathbf{W} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

Recurrent (coupled)
network

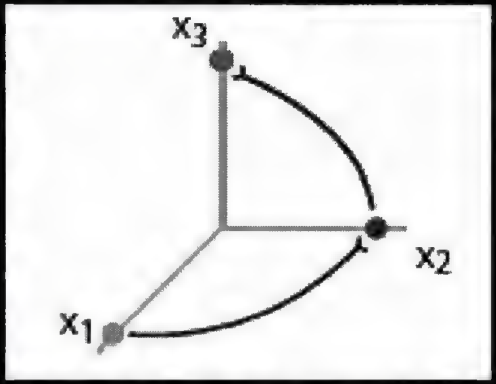


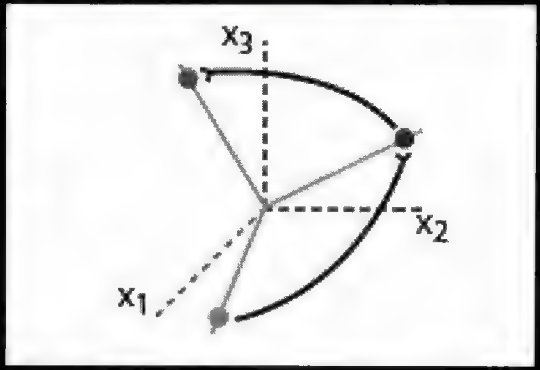
$\mathbf{W}_{recurrent} = \mathbf{R}\mathbf{W}\mathbf{R}^{-1}$

(Schur decomposition)

Map each neuron to
a combination of
neurons by applying
a coordinate rotation
matrix **R**

Geometric picture:

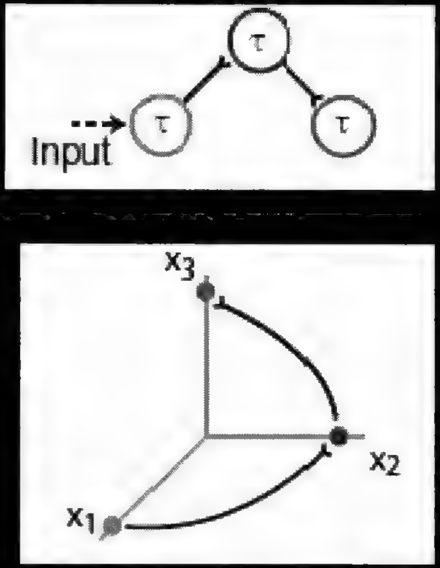




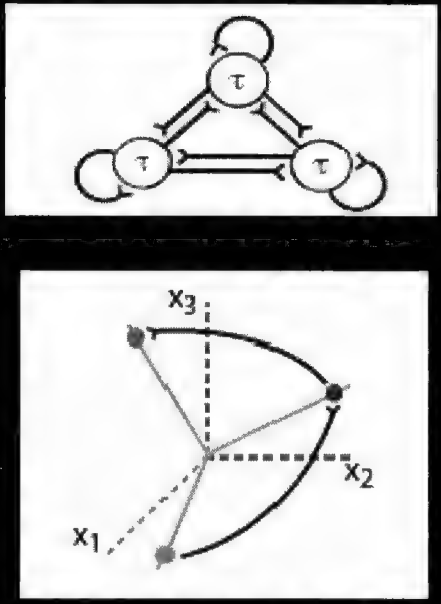
(Math of Schur decomposition: See Goldman, *Neuron*, 2009; Murphy & Miller, *Neuron*, 2009)

Responses of functionally feedforward networks

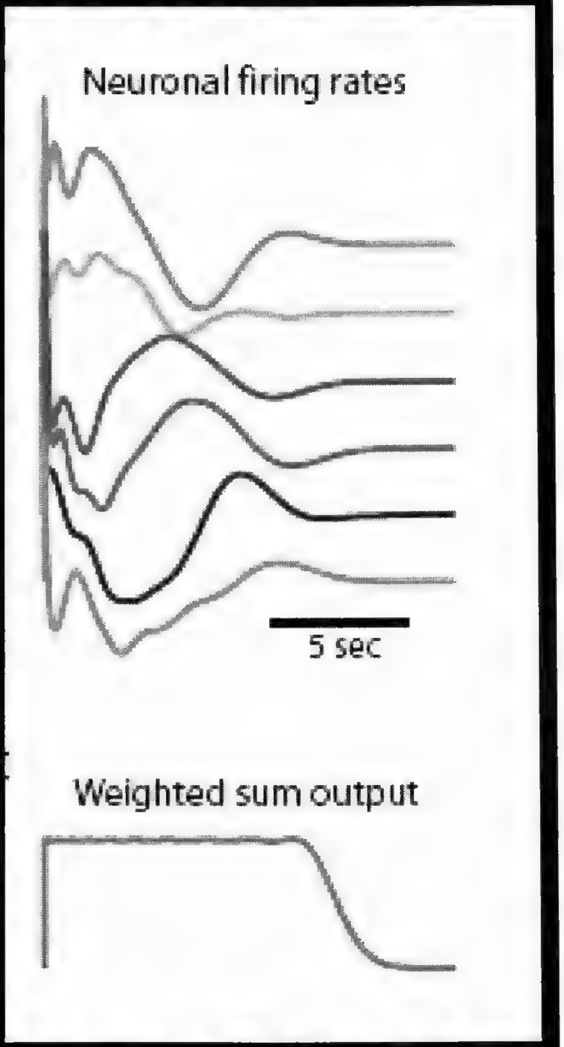
Feedforward
network activity patterns


Input
τ
τ
τ
x_3
x_2
x_1

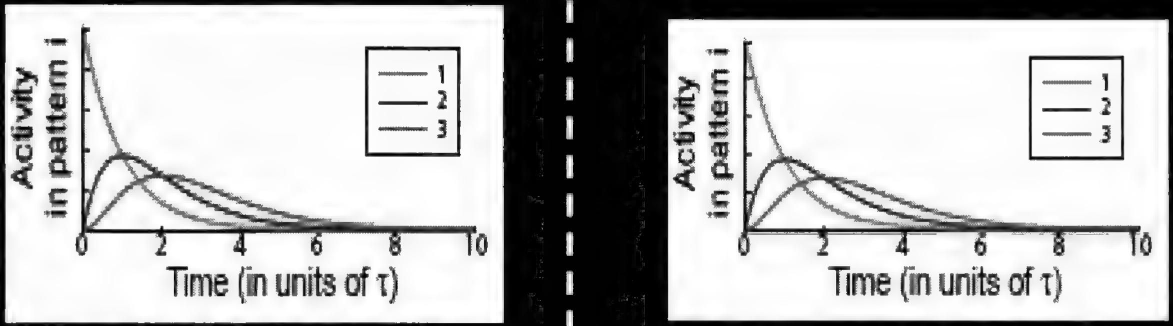
Functionally feedforward
activity patterns…



& neuronal firing rates



Effect of stimulating pattern 1:

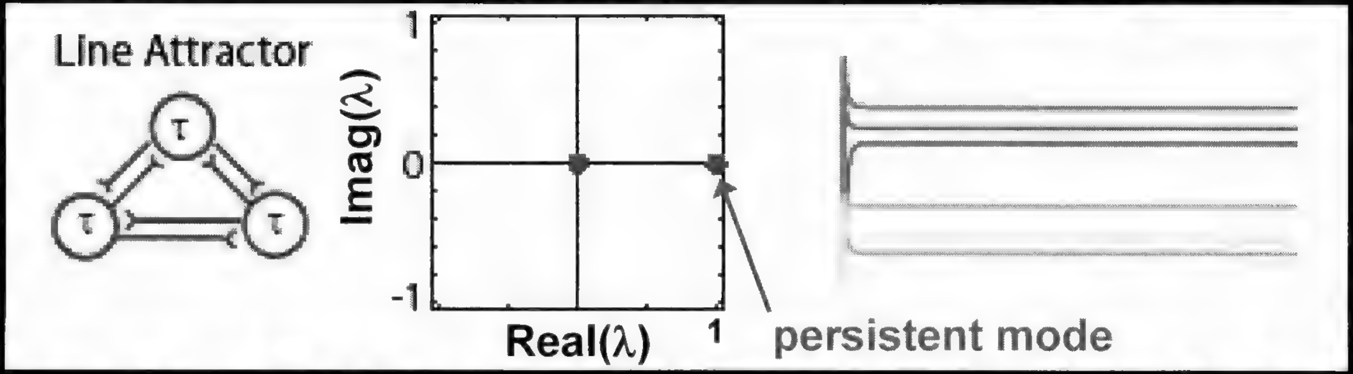


Math Puzzle: Eigenvalue analysis does not predict long time scale of response!

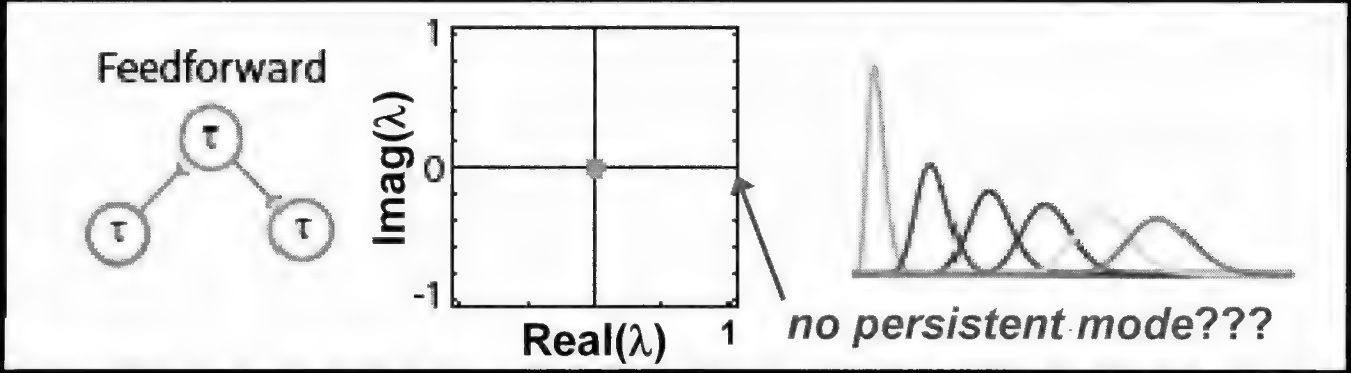
Line attractor networks:

Eigenvalue spectra:

Neuronal responses:


Line Attractor
τ
τ
τ
Imag(λ)
1
0
-1
Real(λ)
1
persistent mode

Feedforward networks:


Feedforward
τ
τ
τ
Imag(λ)
1
0
-1
Real(λ)
1
no persistent mode???

(Goldman, *Neuron*, 2009; see also: Murphy & Miller, *Neuron* 2009; Ganguli & Sompolinsky, PNAS 2008)

Summary

- Short-term memory (~10’s seconds) is maintained by *persistent activity* following the offset of a remembered stimulus

- Possible mechanisms

  1) Tuned positive feedback:

    -Activity perpetuated through positive feedback

    -Dynamics can be described by attractors

  2) Feedforward network (possibly in disguise)

    -*Disadvantage*: Finite memory lifetime ~ # of feedforward stages

    -*Advantage*: Higher-dimensional representation can produce many different temporal response patterns

    -*Math:* Not well-characterized by eigenvalue decomposition; Schur decomposition or pseudospectral analysis better

Acknowledgments

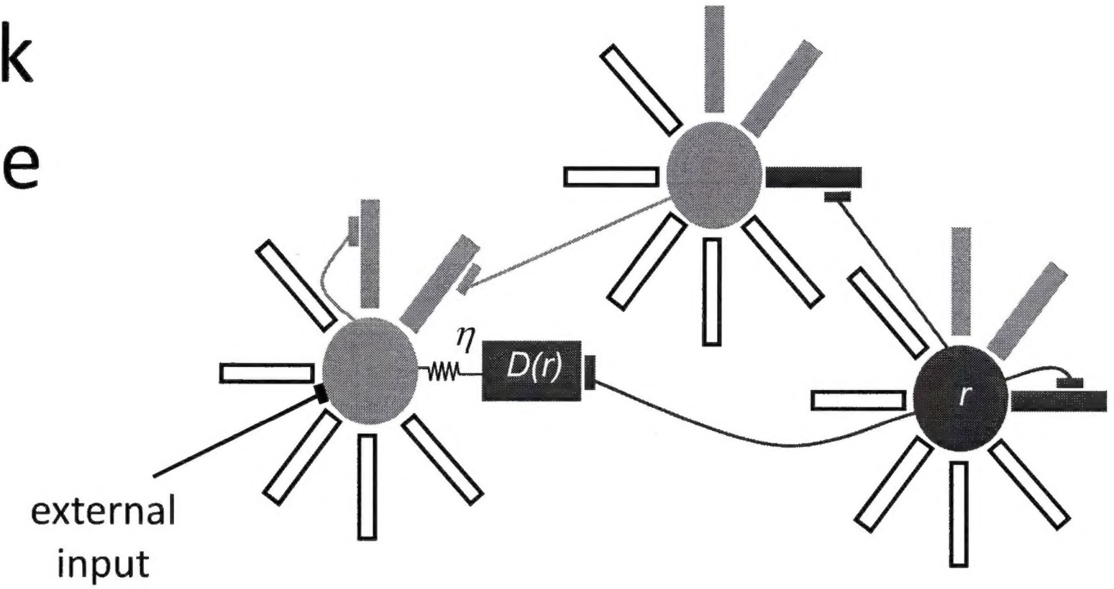
Theory (Goldman lab, UCD)

Itsaso Olasagasti (USZ)
Dimitri Fisher

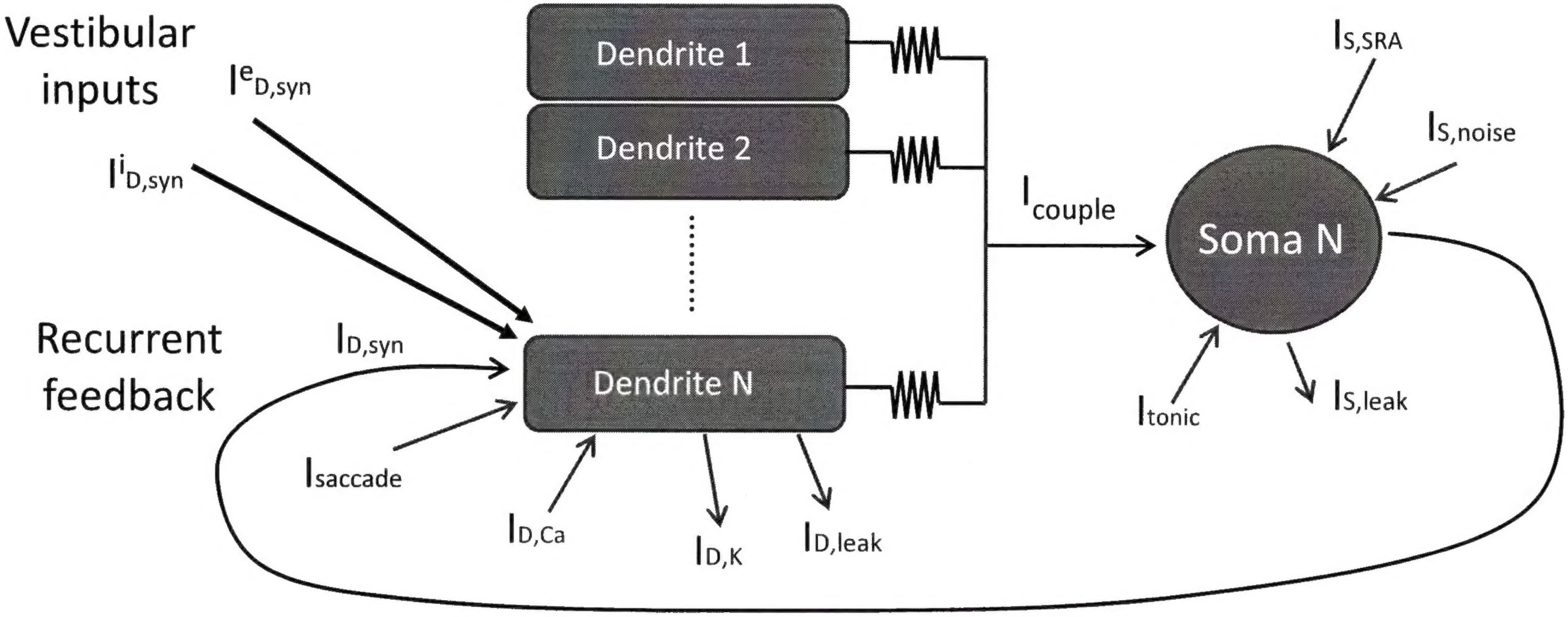
Experiments

David Tank (Princeton Univ.)
Emre Aksay (Cornell Med.)
Guy Major (Cardiff Univ.)
Robert Baker (NYU Medical)

Network structure



η
$D(r)$
r
external input

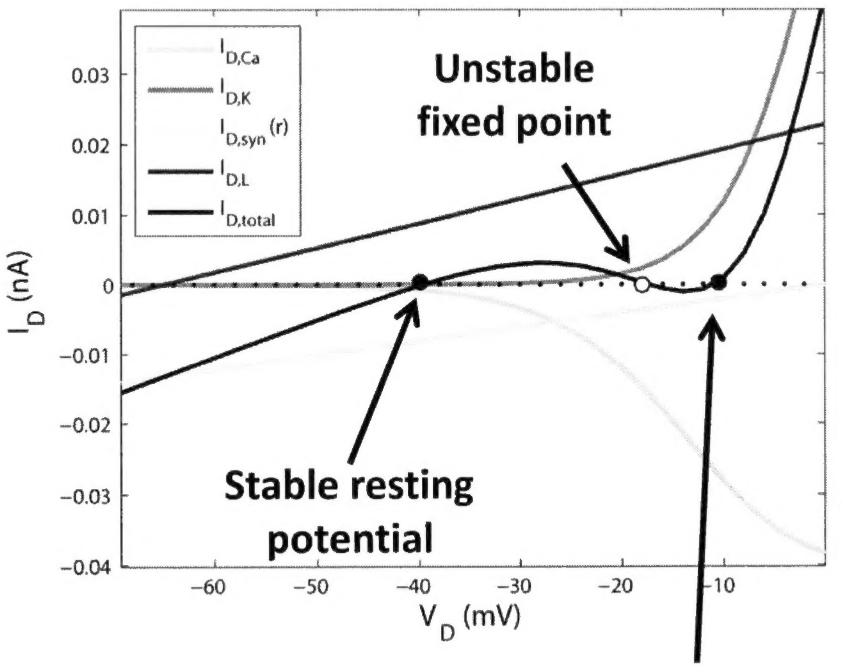


Vestibular inputs
$I^{e}_{D,syn}$
$I^{i}_{D,syn}$
Recurrent feedback
$I_{D,syn}$
$I_{saccade}$
Dendrite 1
Dendrite 2
Dendrite N
$I_{D,Ca}$
$I_{D,K}$
$I_{D,leak}$
I_{couple}
Soma N
$I_{S,SRA}$
$I_{S,noise}$
I_{tonic}
$I_{S,leak}$

Dendritic compartment

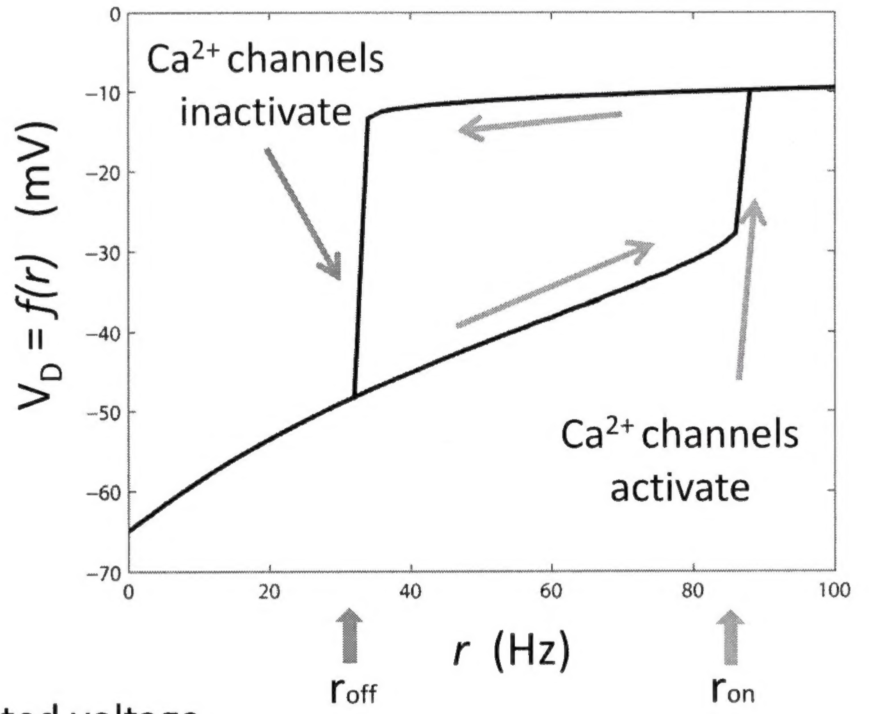
- **Morris-Lecar** formalism, adapted from Booth & Rinzel (1995)
- **Bistability:** two stable voltage levels are maintained by intracellular feedback loops due to voltage-sensitive currents (Ca^{2+}, K^+)
- **Threshold for activation** (and deactivation): presynaptic firing rate must exceed a threshold before it activates a plateau potential

Steady-state current-voltage relation
Input rate = 70 Hz



Plateau potential: elevated voltage triggered by presynaptic firing rate but maintained in its absence

Steady-state voltage-firing rate relation

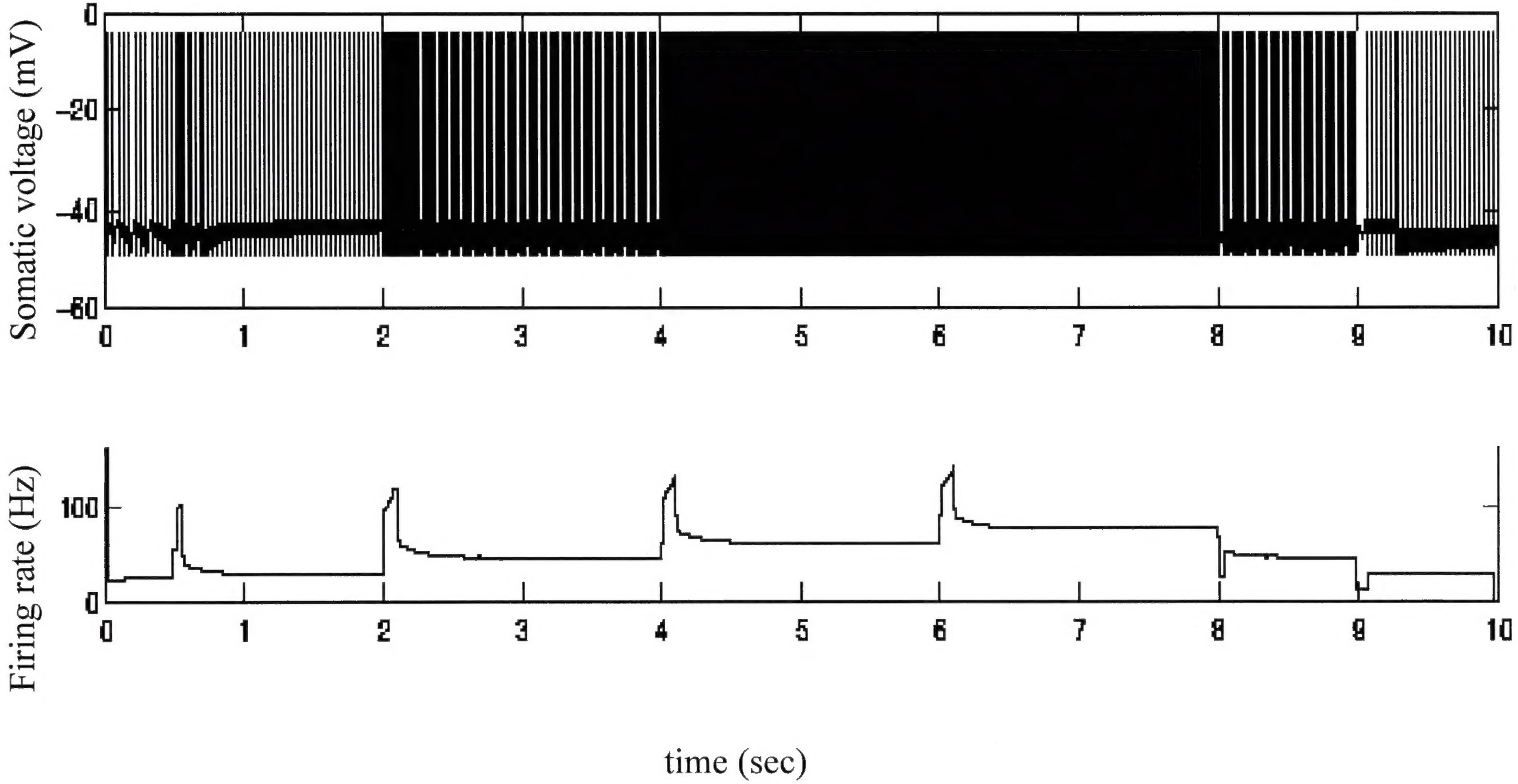


Biophysical Model

Network of 20 recurrently connected neurons, each with:

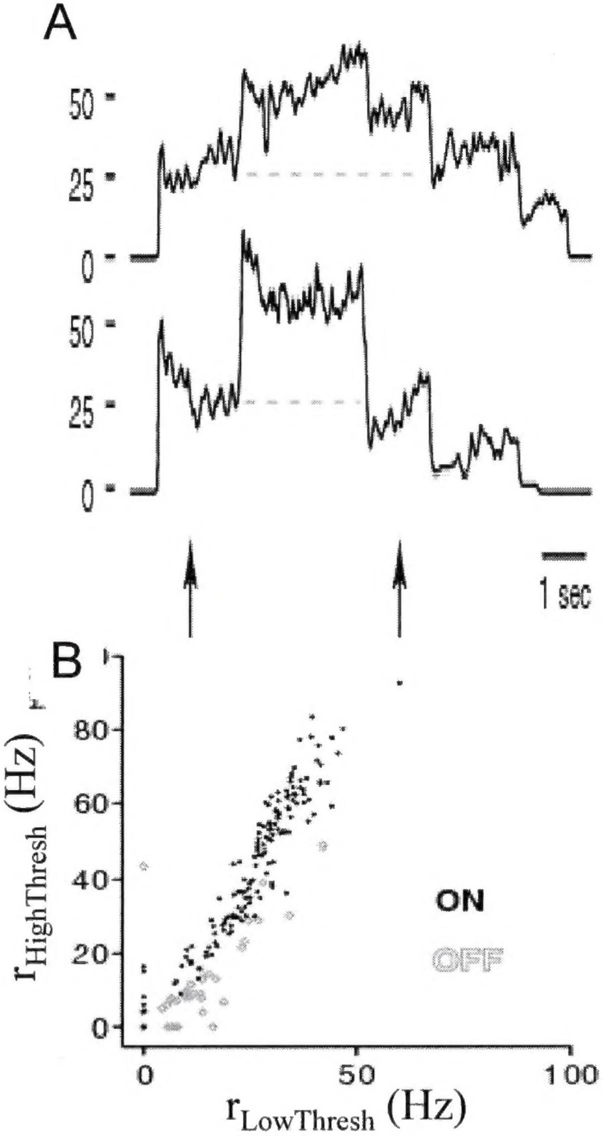
(Simulations by Joseph Levine)

- Spiking model soma
- Calcium plateau mediated bistability
- Soma and dendrites ohmically coupled



Hysteresis in neuronal responses

EXPERIMENT



MODEL